

# **Thèse de Doctorat**

*Spécialité Mathématiques Appliquées*

présentée à

**L'Université de Picardie Jules Verne**

par

**Julien Caron**

pour obtenir le grade de Docteur de l'Université de Picardie Jules Verne

***Restauration en échantillonnage irrégulier  
Théorie et applications aux signaux et images satellitaires***

Soutenue le ....., après avis des rapporteurs, devant le jury d'examen :

<b>M. J.-C. Pesquet, Professeur</b>	<b>Rapporteur</b>
<b>M. F. Malgouyres, Professeur</b>	<b>Rapporteur</b>
<b>M. O. Goubet , Professeur</b>	<b>Examineur</b>
<b>M. B. Rougé , Directeur de Recherches Associé</b>	<b>Examineur</b>
<b>M. J. Fadili, Professeur</b>	<b>Examineur</b>
<b>M. S. Durand, Professeur</b>	<b>Directeur de thèse</b>
<b>M. A. Almansa, Chargé de Recherches</b>	<b>Directeur de thèse</b>



# **Irregular sampling restoration Theory and applications to satellite images and signals**

Julien Caron



# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Modern push-broom satellites . . . . .	5
1.1.1	SPOT5 and PLEIADES satellites . . . . .	5
1.1.2	Acquisition and irregularity in the data . . . . .	6
1.2	Static interferometry . . . . .	8
1.3	Other examples . . . . .	9
1.4	Overview and contributions of the thesis . . . . .	11
<b>2</b>	<b>Mathematics</b>	<b>13</b>
2.1	Fourier Analysis and approximation . . . . .	13
2.1.1	The Fourier transform . . . . .	13
2.1.2	Bandlimited functions . . . . .	14
2.1.3	Stable and interpolation sets, Frame property . . . . .	15
2.1.4	Bandpass functions . . . . .	16
2.1.5	Approximation of bandlimited functions with polynomials . . . . .	17
2.1.5.1	Trigonometric polynomials, Toeplitz operators, convergence . . . . .	17
2.1.5.2	Splines, interpolating splines and convergence theorem . . . . .	19
2.2	Optimization in Hilbert spaces . . . . .	22
2.2.1	Proper functions, Convexity, continuity, semi-continuity, inf-sup stability . . . . .	22
2.2.2	Properties and characterization of convex functions . . . . .	24
2.2.3	Subdifferential and Legendre transform . . . . .	24
2.2.4	Lagrangian, duality gap, Fenchel-Moreau duality . . . . .	26
2.2.5	The proximal operator, examples . . . . .	27
2.2.6	Minimization algorithms and convergence results . . . . .	28
2.2.6.1	The Conjugate Gradient algorithm . . . . .	28
2.2.6.2	Nonsmooth optimization : Forward-Backward, FISTA, TWIST, Nesterov . . . . .	29
<b>3</b>	<b>Estimation of the microvibrations</b>	<b>33</b>
3.1	Exploitation of the disparity map . . . . .	33
3.1.1	Introduction . . . . .	33
3.1.2	State of the art . . . . .	34
3.1.3	Framework of this study . . . . .	35
3.1.4	The MARC2 software package . . . . .	35
3.1.5	Information in the disparity map . . . . .	37
3.2	Detection by $l^1$ regularization . . . . .	39
3.2.1	A nonlinear solver for sparse signals . . . . .	39

3.2.2	Approximation with trigonometric polynomials . . . . .	39
3.2.3	Variational formulations of the problem . . . . .	42
3.2.3.1	General formulation . . . . .	42
3.2.3.2	Stationary case . . . . .	43
3.2.4	Resolution of problem $(\mathcal{P}_1)$ and its variations . . . . .	43
3.2.4.1	Minimization with the FB/ISTA algorithm . . . . .	43
3.2.4.2	Detection results . . . . .	45
3.2.5	Partial correction . . . . .	51
3.2.6	Stationary microvibrations . . . . .	53
3.3	Amplitude estimation . . . . .	55
3.3.1	Iterated Hard-Thresholding Algorithm (IHTA) . . . . .	55
3.3.2	A contrario least-squares . . . . .	59
3.3.3	Conclusion of the estimation step . . . . .	61
3.4	An additional hypothesis for better correction . . . . .	62
3.4.1	Formulation of the piecewise affine hypothesis . . . . .	62
3.4.2	Regularization on the second order derivatives . . . . .	64
3.4.2.1	Choice of an $\ell^1$ -norm . . . . .	64
3.4.2.2	The Hessian operator and related variational problems . . . . .	64
3.4.2.3	Minimization algorithms . . . . .	66
3.4.3	The TV2-denoising problem . . . . .	66
3.4.4	Global minimization . . . . .	69
3.4.4.1	Alternate minimization . . . . .	69
3.4.4.2	Direct minimization schemes . . . . .	71
3.4.4.3	Results . . . . .	73
3.4.4.4	Numerical observations . . . . .	75
3.5	Conclusion . . . . .	78
<b>4</b>	<b>Restauration of irregularly sampled images with splines</b>	<b>79</b>
4.1	Introduction . . . . .	79
4.1.1	Two irregular sampling problems . . . . .	80
4.1.2	State-of-the-art . . . . .	81
4.2	Restauration with splines . . . . .	83
4.2.1	Stable resampling . . . . .	83
4.2.2	Quadratic regularization on the Hessian . . . . .	84
4.2.2.1	Variational formulation . . . . .	84
4.2.2.2	Conjugate gradient and Uzawa's algorithm . . . . .	87
4.2.3	On the deblurring problem . . . . .	89
4.2.3.1	Convolution and projection on spline space . . . . .	89
4.2.3.2	Consistency with Shannon's conditions . . . . .	92
4.2.3.3	Variational formulation . . . . .	93
4.2.3.4	Proximal of TV . . . . .	95
4.2.3.5	The whole algorithm . . . . .	97
4.2.3.6	Experiments . . . . .	99
4.2.3.7	Observations . . . . .	106
4.2.3.8	Influence of some parameters . . . . .	106
4.2.4	Perspectives . . . . .	109
4.2.4.1	Total variation at half-integers . . . . .	109
4.2.4.2	Decomposition of the convolution operator . . . . .	112
4.3	Conclusion . . . . .	114

<b>5</b>	<b>A study of bandpass signals in interferometry</b>	<b>115</b>
5.1	An application of bandpass signals sampling in Static Interferometry . . . . .	115
5.1.1	Presentation of the SIFTI static interferometer . . . . .	115
5.1.2	CNES's state of the art . . . . .	120
5.1.3	Questions . . . . .	121
5.1.4	Contents of this chapter . . . . .	121
5.2	Sampling in $BP(A, B)$ . . . . .	122
5.2.1	Link with Static Interferometry . . . . .	122
5.2.2	Periodic Sampling . . . . .	122
5.2.3	Perturbed Sampling . . . . .	124
5.3	Reconstruction from a small number of irregular samples . . . . .	126
5.3.1	A study of CNES methods . . . . .	126
5.3.2	The general setting : nonharmonic Fourier series . . . . .	128
5.3.2.1	Reconstruction is a projection operator . . . . .	129
5.3.2.2	The bad condition number of the orthogonal projector . . . . .	131
5.3.2.3	Oblique projection . . . . .	134
5.3.3	A generalization of existing methods . . . . .	134
5.3.4	Focus on Interlaced Sampling . . . . .	135
5.3.4.1	Non perturbed case . . . . .	135
5.3.4.2	Perturbed case . . . . .	139
5.4	Link with the dual basis . . . . .	140
5.5	Conclusion . . . . .	143
5.6	Appendix . . . . .	145
5.6.1	Perturbation constant $\frac{\ln(2)}{2\pi B}$ . . . . .	145
5.6.2	Non zero determinant . . . . .	146
5.6.3	Measure of the zero set . . . . .	146
5.6.4	Best reconstruction operator . . . . .	146
5.6.5	Perturbation constant $\frac{\ln(1+1/\kappa)}{2\pi B}$ . . . . .	147
<b>6</b>	<b>Conclusion</b>	<b>149</b>



# Chapter 1

## Introduction

### 1.1 Modern push-broom satellites

#### 1.1.1 SPOT5 and PLEIADES satellites

Push-broom acquisition is a specificity of moving instruments with trajectory close to a rigid translation, such as image scanners for documents digitalization and duplication. This is also the case of several satellites, we describe below two families of push-broom satellites designed at CNES which part of our work is applied to.

SPOT5 is the fifth satellite of the SPOT series, it was launched in May 2002 by Ariane 4 from Kourou. Its major applications are agriculture, 2D and 3D mapping, forest observation, natural disasters evaluation, telecommunications. Its altitude is approximately 822 km and its trajectory is nearly circular. Its agility enables a full coverage of the Earth's surface within 26 days but most areas, except at the equator, can be observed more frequently (up to five days for France latitude).

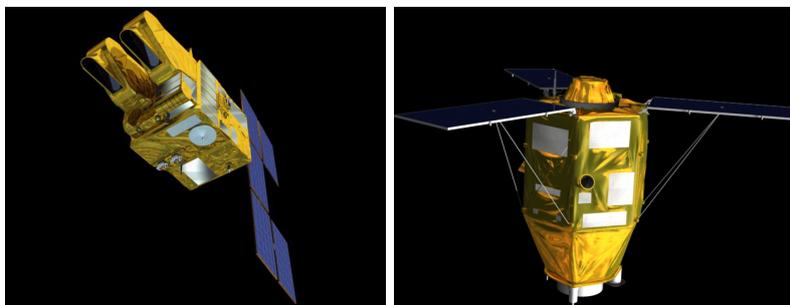


Figure 1.1: SPOT5 (left) and PLEIADES (right) 3D view.

The satellite includes two HRG instruments (High Resolution Graphic), one HRS (High Resolution Stereoscopic) and one VEGETATION instrument (figure 1.2). They have different capabilities in terms of resolution, angle of observation, wavelengths and agility. The VEGETATION instrument is still, has a wide angle and covers a very large swath but has a low resolution (1 km). The HRS instrument is made of two panchromatic sensors with a resolution of 10 meters, they are fixed and the angle between them is  $40^\circ$  in the trajectory direction which allows the acquisition of two views from the same scene with a large tilt. The HRG instrument is made of two identical parts which consist in multi-spectral sensors with 10 meters resolution and a panchromatic sensor with

a 5 meters resolution (at nadir) with agility in the roll angular direction. When the two parts image the same region of the Earth it is possible to obtain a 2.5 meters resolution image (called Supermode).

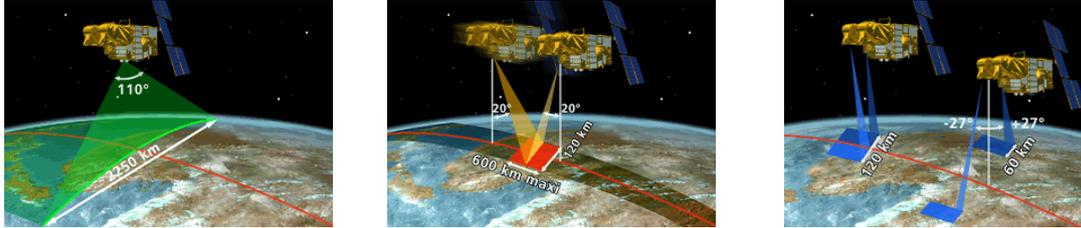


Figure 1.2: SPOT5 instruments (from left to right) : VEGETATION, HRS and HRG.

PLEIADES is the name of two identical satellites, the first one was launched at the end of year 2011 by Soyouz in Kourou, with improved agility and resolution compared to the SPOT series. It is the optical part of the ORFEO program a cooperation between France and Italy and its main applications will be defence, cartography (2D and 3D), geology, geophysics, hydrology, agriculture, forestry and coastal observation. With a 694 km altitude circular trajectory and improved agility, any region of the Earth can be imaged within 24 hours and stereoscopic or tri-stereoscopic will be available with low tilts between the different views. All these possibilities are displayed in figure 1.3.



Figure 1.3: PLEIADES capabilities (from left to right) : adjacent regions, disjoint regions, tri-stereoscopic acquisition.

### 1.1.2 Acquisition and irregularity in the data

From the moment light is emitted by an object to the digital image received from the satellite the signal undergoes many geometrical and numerical transformations as represented in figure 1.4.

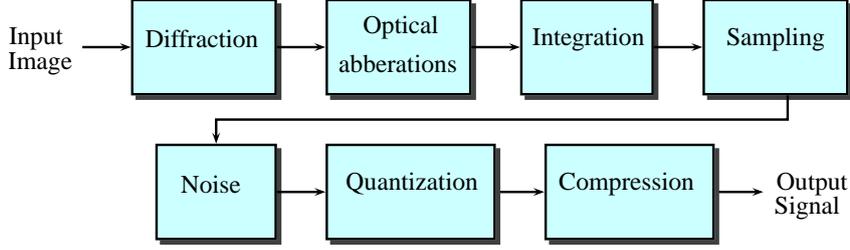


Figure 1.4: Classical modelization of the optical chain in a satellite instrument : the input image is the projection in the focal plane of the light intensity of the scene, it is affected by the diffraction phenomenon as well as other optical aberrations, then integrated and sampled by the detector matrix. Depending on the photon's arrival it can also suffer from different kinds of noise. The energy received by each sensor during the integration is counted, quantized and compressed before transmission.

The push-broom acquisition geometry is slightly different from the classical pin-hole camera model because of the sensors shape and the trajectory of the optical center during acquisition. Furthermore the angular state of the instrument may not be exactly the same during the imaging of a scene, some angular perturbations due to internal mechanisms may occur called microvibrations. As a consequence, the sampling of push-broom images is not exactly a regular square or rectangular grid, in this report we consider the simplified model of an irregular sampling as described in figure 1.5, with the assumption of a small angle  $\gamma$  and no rotation along the  $\vec{X}$  axis (no roll, only pitch).

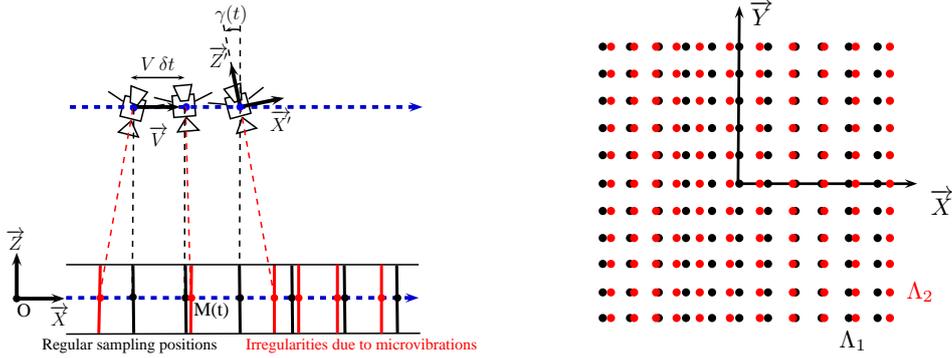


Figure 1.5: Sampling of a push-broom device at nadir with a small pitch  $\gamma(t)$  varying along the trajectory (left) and sampling grid (right). Here the perturbation is assumed horizontal and depends only on the abscissa.

In the SPOT5 or PLEIADES cases, the influence of atmosphere reduces to a simple attenuation and we use the following model image formation model :

$$I(k, l) = (h_{opt} * h_{sens} * h_{mov} * u)(x_{k,l}, y_{k,l}) + \nu_{k,l} \quad (1.1)$$

$I$  is the raster data image,  $h_{opt}$  is the aperture diffraction and optical blurring PSF (Point Spread Function),  $h_{sens}$  is the PSF of a sensor element (generally a square),  $h_{mov}$  is the PSF due to the satellite movement,  $u$  is the input illumination,  $\Lambda = \{(x_{k,l}, y_{k,l})\}$  is the sampling set and  $\nu_{k,l}$  the realization of i-i-d random variables.

The sampling set  $\Lambda$  is generally a small perturbation of the regular grid  $\Lambda_{reg} = \{k(V\delta t), l\delta y\}$  where  $V$  is the satellite tangential speed,  $\delta t$  the time interval between two integrations, and  $\delta y$  the spacing between the sensor elements. We consider in this report two forms of sampling sets  $\Lambda$ , in a first part this set has the form described in figure 1.5 and in a second part it is more general, the perturbations can be two-dimensional and depend on  $x$  and  $y$ .

## 1.2 Static interferometry

Interferometry is the study of signals resulting from an interference phenomenon described in chapter 5. It is used in spectrum reconstruction using a collection of interference measurement also called interferogram whose values are known at regularly spaced locations or on a periodic grid (see figure 1.6 below).

The bandlimitedness of signals is the key point in the design of these sampling sets. In the case of simultaneous acquisitions (static interferometry) the instrument contains stepped mirrors responsible for the optical path differences which form the sampling set. The design of these stepped mirrors has a limited accuracy, this produces irregularities in the sampling grid and reconstruction methods differ from the classical periodic case (figure 1.7).

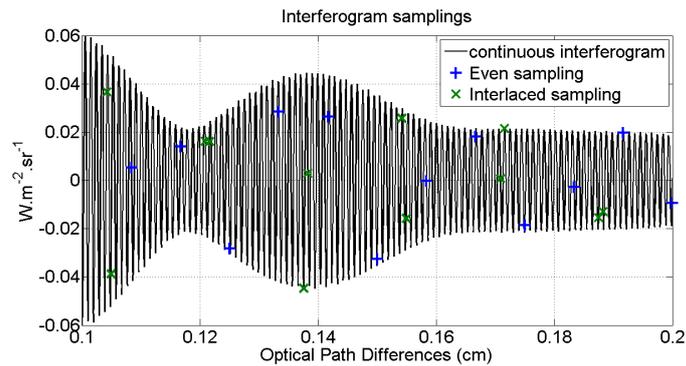


Figure 1.6: In the ideal case the interferogram is sampled on a periodic set, either the regular even sampling (blue) or the interlaced sampling (red) which consists in a series of doublets with intra-doublet spacing  $\epsilon$  and inter-doublet spacing  $dx \gg \epsilon$ .

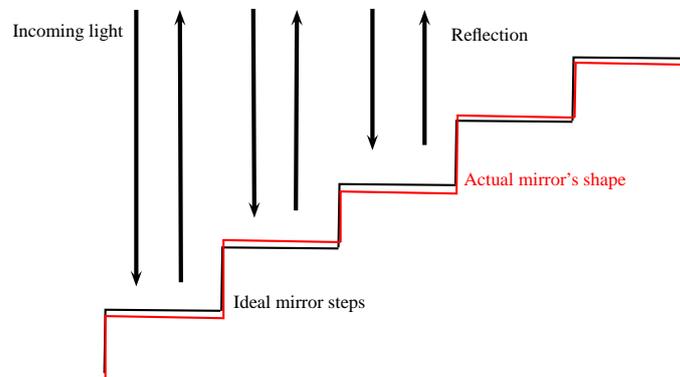


Figure 1.7: Stepped mirror from a static interferometry instrument. Limited accuracy in the steps (red) is responsible for irregular sampling in the interferogram (figure 1.8 below)

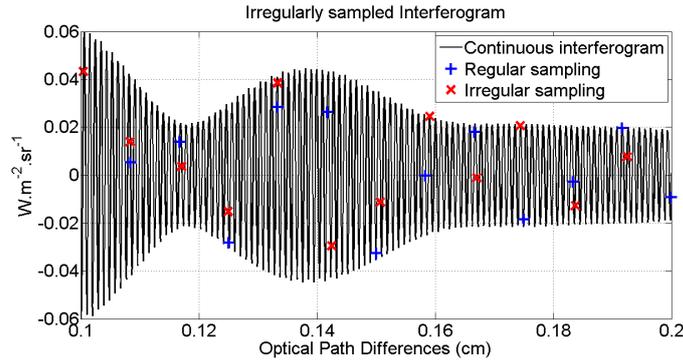


Figure 1.8: The shape of the mirrors steps (figure 1.7) leads to a perturbed sampling of the interferogram. The standard deviation of the perturbation here is typically a tenth of the step size. The resulting optical path differences are spaced by twice the step size and the perturbation is also twice the perturbation on the steps position.

### 1.3 Other examples

The case of images fusion from the same scene but with different angles/positions of the acquisition device is also an irregular sampling problem called super-resolution. Figure 1.9 illustrates the form of the sampling set when a camera has a rigid motion (translation/rotation) between two consecutive images (See Nguyen et al. [71], Vandewalle et al. [91], Rochefort et al. [75] for more details on super-resolution).

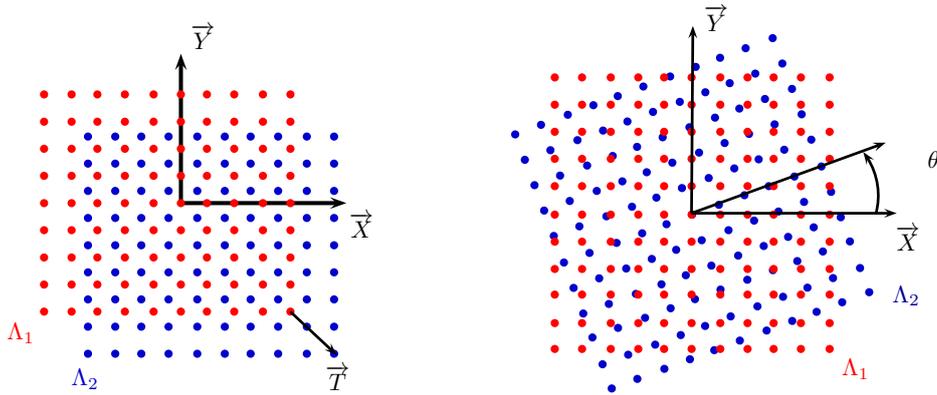


Figure 1.9: Translation and rotation between two sampling grids in the super-resolution problem.

In the two particular cases of figure 1.9 the sampling set is the union of two sets  $\Lambda = \Lambda_1 \cup \Lambda_2$  and tools such as Fourier transform and wavelet transform can be used for these rigid motions, but in the more general case of a non-rigid motion and/or a change in the resolution more general modellings are necessary.

The sampling problem also occurs in some more exotic acquisition devices, for example cameras using fish-eye lenses or spherical/conical mirrors have a much larger angle of view (close to  $180^\circ$ ). The transformation of these omni-directional images into classical geometry images is an irregular sampling problem, although these images are generally exploited in their native spherical geometry [11].

Some geophysical measurements also belong to the irregular sampling field because of the

difficulty in placing sensors at regularly spaced locations (seismic data) or the trajectory of the remote sensing instrument (atmospherical/astronomical data).

In fact irregular sampling is present in nearly every signal measurements if one consider the exact spatial or temporal location of the samples with infinite accuracy. Fortunately the classical mathematical tools are generally sufficient for a very good approximation of the acquisition process as it is the case for most images. Nevertheless it is sometimes necessary to consider more general mathematical tools to efficiently modelize the sampling process, this thesis illustrates several examples from this case.

## 1.4 Overview and contributions of the thesis

This thesis report is divided in three parts, the first two parts address problems from the satellite acquisition field, we first treat the problem of estimation of microvibrations during a satellite flight by exploiting a couple of images from the same scene, and then consider the correction of these images when the sampling grid is slightly irregular as it is the case with push-broom devices. The third part concerns an application to interferometry in which an instrument contains stepped mirrors with irregularities in the steps responsible for the irregular sampling of the interferogram.

The microvibrations estimation from a couple of stereo-rectified images is an ill-posed problem for which only manual estimation exists [78]. The problem was brought to our attention by B. Rougé<sup>1</sup> and the resolution we present here was the opportunity for a fruitful collaboration with L. Moisan<sup>2</sup> and CNES members J. Michel, G. Blanchet. The final step of this collaboration was the delivery of a software package for microvibrations correction. We proposed a modelling of the disparity measurements which takes into account the microvibrations in each image and a detection/correction method based on Basis Pursuit [65]. The algorithms we developed solve variational methods with non-quadratic functions ( $l^1$  norm and  $l^0$  criterion) using tools from convex analysis [30] [8] and nonconvex theory [13]. All these algorithms apply to lacunary disparity measurements as it happens to be the case with many elevation estimation softwares and no manual selection is required. We obtain experimental results showing the benefits of these non-quadratic functions in a very ill-posed problem. When the correlation between the elevation and the microvibrations is too important, these methods partially fail to separate the two components, this is why we propose an improvement based on the piecewise affine behaviour of urban elevation models. This more complicated variational problem can still be handled efficiently with similar tools from convex theory and we obtain a separation of the elevation component and the microvibration component for a class of well segmented data.

The second part concerns the correction of images where the sampling set is slightly irregular as it is the case in the presence of microvibrations. We adopt a functional setting different from previous works of Almansa et al. [4] [46] and show that the spline functions can actually solve this problem with a much lower numerical complexity. We first present some results we obtained with a quadratic regularization on the second order derivatives in the simple resampling case (no deblurring) inspired by works of Arigovindan et al. [5] as an introduction to spline variational methods, part of this work also lead to the delivery of a software package to CNES. Our contribution here is the adaptation of the spline method to the deblurring problem with a fully automatic algorithm using global constraints and Total Variation. We use an approximation of the convolution kernel which is known only in the Fourier domain and solve this problem with a Forward-Backward type algorithm where the convolution is decoupled from the inner proximal loop. The overall algorithm has performances similar to the state-of-the art methods but the computation time is reduced by a factor 10 and up to a factor 20 with low noise levels.

Last part of this thesis is devoted to the study of the general irregular sampling of bandpass signals applied to the static interferometry instrument SIFTI [16] [20] [51] [74]. This work is a collaboration with CNES members D. Jouglot and C. Pierrangelo together with S. Ladjal<sup>3</sup> who brought to our attention a certain number of numerical and theoretical questions about the reconstruction of light spectra from irregularly sampled interferograms. After a description of the

---

<sup>1</sup>CESBIO, CNES Toulouse

<sup>2</sup>University Paris Descartes

<sup>3</sup>Telecom ParisTech

instrument and its specifications, we study the state-of-the-art algorithms and propose a generalization of these methods with correction of some numerical approximations. This more general approach is based on nonharmonic Fourier series, thanks to this approach we can give the exact expression of the result for different reconstruction families and justify the differences observed in the state-of-the-art methods. We then study the stability of the reconstruction operator in terms of conditioning when the sampling set is perturbed by random i.-i.-d. Gaussian variables. We show several theorems on the allowed maximal perturbation in bandpass spaces when the sampling set is infinite, although our experiments with finitely many samples tend to show that random perturbations with amplitudes larger than the theoretical bound still allow correct reconstruction. At last we study in details the interlaced sampling and its stability properties compared to regular sampling and show experimentally that this sampling is indeed more stable thanks to a well conditioned operator whose singular values are almost always bounded by below.

# Chapter 2

## Mathematics

This chapter concerns the mathematical tools used in this thesis, they belong to two very important fields of mathematics, namely Harmonic Analysis and Optimization. They are used in the resolution of problems arising in satellite imaging, depth estimation and interferometry. We refer the reader to Katznelson [55] in Harmonic Analysis, Young [95] in Nonharmonic Analysis, Mallat [65] in Signal Processing and Wavelet theory. Concerning Optimization theory, we use mainly results from Rockafellar [76], Bertsekas [10] and Combettes [29].

### 2.1 Fourier Analysis and approximation

#### 2.1.1 The Fourier transform

Fourier transform is a very general transform which can be defined on any locally compact (topological) commutative group. In Signal Processing it is generally sufficient to know its properties on the classical additive groups  $\mathbb{R}$ ,  $\mathbb{Z}$ ,  $\mathbb{Z}/n\mathbb{Z}$ ,  $\mathbb{R}/w\mathbb{Z}$  and their higher dimension versions. We use the Fourier transform defined on  $L^1(\mathbb{R}^d)$  by

$$\mathcal{F} : f \in L^1(\mathbb{R}^d) \rightarrow \hat{f} : \xi \rightarrow \int_{\mathbb{R}^d} f(x) e^{i \langle x | \xi \rangle} dx \quad (2.1)$$

but also on  $L^2(\mathbb{R}^d)$  as a limit of this integral. We remind Parseval's and Plancherel's formulas and the derivation/multiplication properties.

$$\langle f, g \rangle_{L^2(\mathbb{R}^d)} = \frac{1}{(2\pi)^d} \langle \hat{f}, \hat{g} \rangle_{L^2(\mathbb{R}^d)} \quad \text{for all } f, g \in L^2(\mathbb{R}^d) \quad (2.2)$$

$$\|f\|_{L^2(\mathbb{R}^d)}^2 = \frac{1}{(2\pi)^d} \|\hat{f}\|_{L^2(\mathbb{R}^d)}^2 \quad \text{for all } f \in L^2(\mathbb{R}^d) \quad (2.3)$$

$$\mathcal{F} \left( \frac{\partial f}{\partial x_k} \right) = i \xi_k \hat{f} \quad \text{if } f, \frac{\partial f}{\partial x_k} \in L^1(\mathbb{R}^d) \quad (2.4)$$

$$\mathcal{F} (x_k f) = -i \frac{\partial \hat{f}}{\partial \xi_k} \quad \text{if } f, x_k f \in L^1(\mathbb{R}^d) \quad (2.5)$$

Fourier transform is linear, injective, continuous  $L^2 - L^2$ ,  $L^1 - L^\infty$  and with our definition it verifies

$$\mathcal{F} (\mathcal{F}(f)) = (2\pi)^d \check{f} \quad \text{for all } f \in L^2(\mathbb{R}^d) \quad (2.6)$$

with  $\check{f}(x) : x \rightarrow f(-x)$ . The convolution product on  $L^1(\mathbb{R}^d)$  defined by

$$(f * g)(x) = \iint_{\mathbb{R}^d} f(t) g(x - t) dt \quad (2.7)$$

has Fourier transform

$$\mathcal{F}(f * g)(\cdot) = \hat{f}(\cdot) \hat{g}(\cdot) \quad (2.8)$$

On  $\ell^2(\mathbb{Z})$  we use the following definition

$$\mathcal{F}((c_k)_{k \in \mathbb{Z}})(\xi) = \sum_{k \in \mathbb{Z}} c_k e^{i k 2\pi \xi} \quad (2.9)$$

which maps  $\ell^2(\mathbb{Z})$  to  $L^2(\mathbb{R}/\mathbb{Z})$  (one-to-one). The inverse transform is the Fourier coefficients application

$$f \in L^2(\mathbb{R}/\mathbb{Z}) \rightarrow (c_k)_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \quad \text{with} \quad c_k = \int_0^1 f(t) e^{-i k 2\pi t} dt \quad (2.10)$$

These two transforms are linked by the very useful Poisson's formula

**Proposition 1 (Poisson's formula)** *Let  $f \in L^2(\mathbb{R})$  such that*

$$|f(x)| \leq \frac{C}{(1 + |x|)^\alpha} \quad |\hat{f}(\xi)| \leq \frac{C}{(1 + |\xi|)^\alpha} \quad \text{for all } x, \xi \in \mathbb{R} \quad (2.11)$$

with  $C > 0$  and  $\alpha > 1$ . Let  $\delta > 0$ , then

$$\sum_{k \in \mathbb{Z}} f(k\delta) e^{i k 2\pi \xi} = \frac{1}{\delta} \sum_{l \in \mathbb{Z}} \hat{f}\left(\frac{2\pi}{\delta}(\xi + l)\right)$$

This formula is crucial in many Signal Processing applications, in particular it provides an interpretation in the Fourier domain of the possible information loss during sampling procedure. In dimension one, the Fourier transform is periodized with step  $\frac{2\pi}{\delta}$ , in higher dimensions the shifts depend on the dual lattice of the sampling grid.

### 2.1.2 Bandlimited functions

Bandlimited functions are functions in  $L^2(\mathbb{R})$  whose Fourier transform is supported in a compact subset of  $\mathbb{R}$ .

**Definition 1 (Bandlimited spaces)** *Let  $B \geq 0$ , we define*

$$BF(B) = \left\{ f \in L^2(\mathbb{R}) \quad \text{s.t.} \quad \text{supp}(\hat{f}) \subset [-B, B] \right\}$$

Poisson's formula holds a.e. if  $f \in BF(B)$  without the bound assumptions 2.11, as a consequence we have the famous Shannon's theorem.

**Theorem 1 (Shannon-Whittaker-Kotel'nikov theorem)** *Let  $B > 0$ . For any function  $f$  in  $BF(B)$  we have*

$$f(\cdot) = \sum_{k \in \mathbb{Z}} f\left(k \frac{\pi}{B}\right) \text{sinc}\left(B\left(\cdot - k \frac{\pi}{B}\right)\right)$$

$$\|f\|_{L^2(\mathbb{R})}^2 = \sum_{k \in \mathbb{Z}} \left| f\left(k \frac{\pi}{B}\right) \right|^2$$

with  $\text{sinc}(x) = \frac{\sin(x)}{(x)}$  if  $x \neq 0$  and  $\text{sinc}(0) = 1$ . These equalities are linked to the notions of orthonormal basis and tight frames in Hilbert spaces.

**Definition 2 (Frame)** Let  $\mathcal{H}$  be a Hilbert space and  $\mathcal{E} = (e_k)_{k \in \mathbb{Z}}$  a family of vector of  $\mathcal{H}$ .  $\mathcal{E}$  is called a frame if there exists  $0 < A < B$  such that for all  $f \in \mathcal{H}$

$$A \|f\|_{\mathcal{H}}^2 \leq \sum_{k \in \mathbb{Z}} |\langle f, e_k \rangle_{\mathcal{H}}|^2 \leq B \|f\|_{\mathcal{H}}^2$$

In finite dimension, any complete family is a frame but this is not the case in infinite dimension. When  $A = B$  in the inequality above, the frame is called a *tight frame*, it is very similar to an orthonormal basis but it is not necessarily independent. In fact frames that fail to be a frame by the removal of one element are *Riesz bases* [95].

**Definition 3 (Riesz basis)** Let  $\mathcal{E} = \{e_k\}_{k \in \mathbb{Z}}$  be a family of vector in a Hilbert space  $\mathcal{H}$ .  $\mathcal{E}$  is called a Riesz basis of  $\mathcal{H}$  if there exists  $0 < A < B$  such that

- For all  $N \in \mathbb{N}$  and all finite sequence  $c = (c_k)_{k=-N..N}$

$$A \sum_{k=-N}^N |c_k|^2 \leq \left\| \sum_{k=0}^N c_k e_k \right\|^2 \leq B \sum_{k=0}^N |c_k|^2$$

- The family  $\mathcal{E}$  is complete in  $\mathcal{H}$

$$\overline{\bigcup_{N \in \mathbb{N}} \text{vect}_{k=-N..N} \{e_k\}} = \mathcal{H}$$

Shannon's theorem states that the family  $\{\text{sinc}(B(\cdot - k\frac{\pi}{B}))\}_{k \in \mathbb{Z}}$  is a tight frame of  $BF(B)$  and gives its reconstruction formula. Interpolation and reconstruction in Hilbert spaces often use this frame property, as detailed in [95].

### 2.1.3 Stable and interpolation sets, Frame property

Two key notions in Hilbert spaces have arisen from sampling theory in Hilbert spaces, stable sampling sets and interpolation sampling sets which reflect the existence, uniqueness and stability problematics of signal reconstruction.

**Definition 4 (Stable sampling set)** Let  $\mathcal{H}$  be a functional Hilbert space on  $\mathbb{R}^d$ ,  $\Lambda = \{z_k\}_{k \in \mathbb{Z}} \subset \mathbb{R}^d$  is called a stable sampling set for  $\mathcal{H}$  if there exists  $A > 0$  such that for every function  $f \in \mathcal{H}$

$$\sum_{k \in \mathbb{Z}} |f(z_k)|^2 < \infty \quad \text{and} \quad \|f\|_{\mathcal{H}}^2 \leq A \sum_{k \in \mathbb{Z}} |f(z_k)|^2$$

**Definition 5 (Interpolation set)** Let  $\mathcal{H}$  be a functional Hilbert space on  $\mathbb{R}^d$ ,  $\Lambda = \{z_k\}_{k \in \mathbb{Z}} \subset \mathbb{R}^d$  is called an interpolation set for  $\mathcal{H}$  if

$$\text{for all } (c_k)_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \text{ there exists } f \in \mathcal{H} \text{ s.t. } f(z_k) = c_k \quad \forall k \in \mathbb{Z}$$

A famous result in approximation theory of bandlimited signals is Landau's theorem [58] which gives necessary conditions on the sampling density (definitions below) to have a stable sampling set and/or of interpolation sets.

**Theorem 2 (Landau [58])** Let  $\Lambda = \{z_k\}_{k \in \mathbb{Z}} \subset \mathbb{R}$ ,  $S \subset \mathbb{R}^d$  a measurable set and  $\mathcal{B}(S)$  the set of functions in  $L^2(\mathbb{R}^d)$  with Fourier transform supported in  $S$ .

If  $\Lambda$  is an interpolation set of  $\mathcal{B}(S)$  then

$$D^+(\Lambda) \leq \frac{\mu(S)}{(2\pi)^d}$$

If  $\Lambda$  is a stable sampling set of  $\mathcal{B}(S)$  and  $S$  is bounded then

$$D^-(\Lambda) \geq \frac{\mu(S)}{(2\pi)^d}$$

with

$$D^+(\Lambda) = \overline{\lim}_{r \rightarrow +\infty} \frac{1}{r} |\Lambda \cap B(0, r)| \quad D^-(\Lambda) = \underline{\lim}_{r \rightarrow +\infty} \frac{1}{r} |\Lambda \cap B(0, r)|$$

In fact Young [95] proved that the stable sampling property in a functional RKHS (*Reproducing Kernel Hilbert Space*) is equivalent to the frame property of the family of vectors  $\{g_k\}_{k \in \mathbb{Z}}$  associated to the continuous linear forms  $f \in \mathcal{H} : \rightarrow f(z_k)$ , that is

$$f(z_k) = \langle f, g_k \rangle \quad \text{for all } f \in \mathcal{H} \quad \text{and } z_k \in \Lambda$$

Indeed bandlimited spaces are RKHS spaces since for all  $f \in BF(B)$  one has

$$\begin{aligned} f(x) &= \frac{1}{(2\pi)^d} \left\langle \hat{f}, \mathbb{1}_{[-B, B]^d}(\cdot) e^{-ix \cdot} \right\rangle_{L^2(\mathbb{R}^d)} \\ &= \left\langle f, \mathcal{F}^{-1}(\mathbb{1}_{[-B, B]^d}(\cdot) e^{-ix \cdot}) \right\rangle_{\mathcal{H}} \quad \text{for all } x \in \mathbb{R}^d \end{aligned}$$

The one-dimensional case is more specific as bandlimited functions benefit from the analytical theory (see Paley-Wiener [73]) in particular bandlimited functions on  $\mathbb{R}$  are completely characterized by their growth rate on  $\mathbb{C}$ . The  $\mathbb{C}$ -extension of a bandlimited function  $f \in BF(B)$  has expression :

$$f(z) = \int_{-B}^B \hat{f}(\xi) e^{iz\xi} d\xi \quad (2.12)$$

**Theorem 3 (Paley-Wiener [73] p. 13)** Let  $f$  be an entire function on  $\mathbb{C}$ , it is the Fourier transform of a function  $g \in L^2(\mathbb{R})$  (formula 2.12) with support in  $[-B, B]$  if and only if  $f$  is square integrable on the real axis and there exists  $C > 0$  such that

$$|f(z)| \leq C e^{B|z|} \quad \text{for all } z \in \mathbb{C}$$

Bandlimited functions have thus an analytic extension to the whole complex space and can be developed in Taylor series :

$$f(x) = \sum_{l \in \mathbb{N}} f^{(l)}(x_0) \frac{(x - x_0)^l}{l!} \quad \text{for all } x, x_0 \in \mathbb{R} \quad \text{and } f \in BF(B) \quad (2.13)$$

### 2.1.4 Bandpass functions

In chapter 5 we study a particular form of bandlimited signals with Fourier support in a union of intervals, in particular signals with Fourier support in two symmetric intervals.

**Definition 6 (Bandpass signals)** Let  $0 < A < B$ , we denote by  $BP(A, B)$  the set of functions in  $L^2(\mathbb{R})$  with Fourier transform supported in  $\Omega = [-B, -A] \cup [A, B]$ .

$$BP(A, B) = \left\{ f \in L^2(\mathbb{R}) \quad \text{s.t.} \quad \text{supp}(\hat{f}) \subset [-B, -A] \cup [A, B] \right\} \subset BF(B)$$

These signals benefit from the generalized Shannon's theorem, as more flexibility is allowed in the frequency periodization, and a frame reconstruction formula.

**Theorem 4 (Generalized Shannon's theorem)** Let  $0 < A < B$  and  $dx > 0$ . The set  $\Lambda = \{k dx\}_{k \in \mathbb{Z}}$  is a stable sampling set of  $BP(A, B)$  if and only if  $dx$  belongs to one of the intervals

$$I_k = \left[ \frac{k\pi}{A}, \frac{(k+1)\pi}{B} \right] \quad \text{with} \quad 0 \leq k \leq \left\lfloor \frac{A}{B-A} \right\rfloor, \quad k \in \mathbb{N}$$

In this case, for any  $f \in BP(A, B)$

$$\begin{aligned} f(\cdot) &= \mathcal{F}^{-1} \left( \left( \sum_{l \in \mathbb{Z}} \hat{f}(\xi + l \frac{2\pi}{dx}) \right) \mathbb{1}_{[-B, -A] \cup [A, B]} \right) (\cdot) \\ &= \sum_{k \in \mathbb{Z}} f(k dx) \frac{(B-A) dx}{\pi} \cos \left( \frac{A+B}{2} (\cdot - k dx) \right) \text{sinc} \left( \frac{B-A}{2} (\cdot - k dx) \right) \end{aligned} \quad (2.14)$$

with similar energy estimates as Shannon's theorem because of the special interactions between intervals  $[-B, -A]$  and  $[A, B]$  in the Fourier periodization when  $dx$  belongs to one of these intervals.

## 2.1.5 Approximation of bandlimited functions with polynomials

### 2.1.5.1 Trigonometric polynomials, Toeplitz operators, convergence

Trigonometric polynomials are closely related to the Fourier transform on the torus  $\mathbb{R}/\mathbb{Z}$  and discrete Fourier transform on  $\mathbb{Z}/n\mathbb{Z}$ . A trigonometric polynomial with degree  $d$  and period 1 is a function  $P : \mathbb{R} \rightarrow \mathbb{C}$  of the form

$$P(t) = \sum_{k=-d}^d p_k e^{i k 2\pi t} \quad t \in \mathbb{R} \quad (2.15)$$

with  $(p_k)_{-d \leq k \leq d}$  its coefficients. The family  $\{e^{i k 2\pi \cdot}\}_{k \in \mathbb{Z}}$  is an orthonormal basis of  $L^2(\mathbb{R}/\mathbb{Z})$  (also denoted by  $L^2([0, 1])$ ) and for every function in this space we have the Parseval and Plancherel formulas :

$$\begin{aligned} \langle f, g \rangle &= \int_0^1 f(t) \overline{g(t)} dt = \sum_{k \in \mathbb{Z}} f_k \overline{g_k} \\ \|f\|^2 &= \int_0^1 |f(t)|^2 dt = \sum_{k \in \mathbb{Z}} |f_k|^2 \end{aligned}$$

and convolution identity :

$$\begin{aligned} (f * g)(\cdot) &= \int_0^1 f(t) g(\cdot - t) dt \\ c_k(f * g) &= f_k g_k \end{aligned} \quad (2.16)$$

with  $(f_k)_{k \in \mathbb{Z}}, (g_k)_{k \in \mathbb{Z}}$  the Fourier coefficients of  $f$  and  $g$  and  $c_k(\cdot)$  the Fourier coefficient application from (2.10).

Thanks to the finiteness of their Fourier coefficients, trigonometric polynomials benefit from the discrete Fourier transform on  $\mathbb{Z}/n\mathbb{Z}$  and its fast implementation.

$$(p_k)_{-d \leq k \leq d} = \mathcal{F}(P) = \frac{1}{2d+1} \sum_{-d \leq l \leq d} f\left(\frac{l}{2d+1}\right) e^{i k 2\pi \frac{l}{2d+1}}$$

Convolution on the space  $\mathcal{B}_d$  of 1-periodic trigonometric polynomials with degree at most  $d$  admits a neutral element called the Dirichlet kernel with formula

$$\begin{aligned} D_d(t) &= \frac{1}{2d+1} \sum_{-d \leq k \leq d} e^{i k 2\pi t} \\ &= \frac{\sin((2d+1)\pi t)}{(2d+1) \sin(\pi t)} \end{aligned}$$

and for all  $P$  in  $\mathcal{B}_d$  we have  $P(\cdot) = \sum_{k=0}^{2d} P\left(\frac{k}{2d+1}\right) D_d\left(\cdot - \frac{k}{2d+1}\right)$ .

Sampling problems in the space  $\mathcal{B}_d$  would not be accessible without the nice following property.

**Proposition 2 (Groechenig et al. [48])** *Let  $\Lambda = \{t_k\}_{k=1..N} \subset [0, 1[$  and  $P \in \mathcal{B}_d$ , the vector  $(P(t_k))_{k=1..2d+1}$  rewrites as a linear combination of the polynomial coefficients*

$$(P(t_k))_{k=1..2d+1} = S_\Lambda(p_l)$$

where  $S_\Lambda$  has Vandermonde structure. Furthermore the product  $S_\Lambda^* S_\Lambda$  has Toeplitz structure

$$(S_\Lambda^* S_\Lambda)_{j,l} = \sum_{k=1..N} e^{-i(j-l)2\pi x_k}$$

This allows fast matrix-vector computation and is completed in works from Groechenig et al. [48] by a preconditioning.

**Proposition 3 (Preconditioned frame operator)** *Let  $\Lambda = \{t_k\}_{k=1..N} \subset [0, 1[$  be a non-decreasing sequence with maximal gap  $0 < \delta = \max_{1 \leq k \leq N} \{t_{k+1} - t_k\} < \frac{1}{2d}$ . Then*

$$(1 - 2\delta M)^2 \|p\|_2^2 \leq \sum_{k=1}^N |P(t_k)|^2 \frac{t_{k+1} - t_{k-1}}{2} \leq (1 + 2\delta M)^2 \|p\|_2^2$$

for all  $P \in \mathcal{B}_d$ . (with convention  $t_0 = t_N - 1$  and  $t_{N+1} = t_1 + 1$ )

A two-dimensional version of this preconditioning is used in [46] using Voronoi areas.

Trigonometric polynomials are often used to approximate bandlimited functions, indeed the least-squares approximation by trigonometric polynomials with equal period and number of coefficients (denoted  $\mathcal{B}_M(2M+1)$ ) converges uniformly on every compact set.

**Theorem 5 (Convergence theorem)** *Let  $\Lambda = \{t_k\}_{k \in \mathbb{Z}}$  be a non-decreasing sequence such that  $\sup_{k \in \mathbb{Z}} |t_{k+1} - t_k| = \delta < 1$  and  $f \in BF(\pi)$ . Let  $P_M$  denote the unique solution of the least-squares problem*

$$\min_{P_M \in \mathcal{B}_M(2M+1)} \sum_{|t_k| < M + \frac{1}{2}} |P_M(t_k) - f(t_k)|^2 \frac{t_{k+1} - t_k}{2}$$

Then

$$\lim_{M \rightarrow +\infty} \int_{-M-\frac{1}{2}}^{M+\frac{1}{2}} |f^{(l)}(t) - P_M^{(l)}(t)| dt = 0$$

for all derivative order  $l \in \mathbb{N}$ .

### 2.1.5.2 Splines, interpolating splines and convergence theorem

Splines are piecewise polynomial functions often used in approximation. Given a nondecreasing sequence  $(t_k)_{1 \leq k \leq N}$  of points called knots, they have a polynomial expression on each segment  $[t_k, t_{k+1}]$  with degree  $d$  and  $d-1$  global regularity. *B-splines* are compactly supported splines with minimal support in  $[t_k, t_{k+d+1}]$  and form a basis of the spline space of order  $d$ . In this thesis we consider regularly spaced knots at integer location (the common framework in image processing) and approximate the image with tensorial B-splines as defined by Unser et al. [89] [90] :

**Definition 7 (B-splines)** We call regular B-spline of order  $d \in \mathbb{N}$  the function  $\beta_d$  defined by iterated convolutions :

$$\begin{aligned} \beta_d(x) &= 1_{[-\frac{1}{2}, \frac{1}{2}[}^{*(d+1)}(x) \\ &= \underbrace{1_{[-\frac{1}{2}, \frac{1}{2}[} * \dots * 1_{[-\frac{1}{2}, \frac{1}{2}[} }_{(d+1) \text{ terms}}(x) \end{aligned}$$

The  $\beta_d$  functions are symmetric, supported in  $[-\frac{d+1}{2}, \frac{d+1}{2}]$ , have a  $\mathcal{C}^{(d-1)}$  regularity as shown in figure 2.1. These functions have an explicit spatial expression :

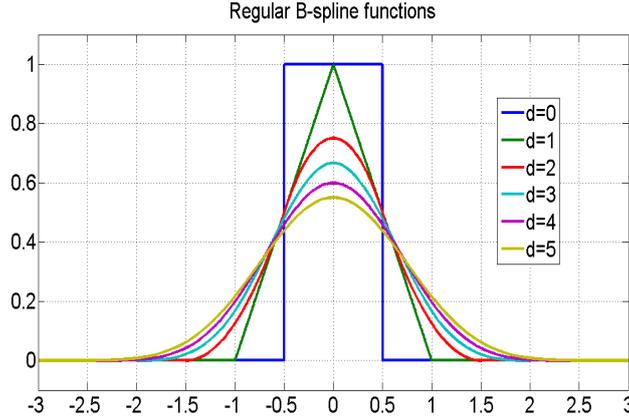


Figure 2.1: Regular B-splines from order 0 to 5.

$$\beta_d(x) = \frac{1}{d!} \sum_{k=0}^{d+1} (-1)^k \binom{d+1}{k} \left(x + \frac{d+1}{2} - k\right)_+^d \quad (2.17)$$

where  $(\cdot)_+$  denotes the positive part. Another very useful formula is a particular case of De Boor's recurrence formula in the case of uniform knots.

$$\beta_d(x) = \frac{1}{d} \left[ \left(\frac{d+1}{2} + x\right) \beta_{d-1}\left(x + \frac{1}{2}\right) + \left(\frac{d+1}{2} - x\right) \beta_{d-1}\left(x - \frac{1}{2}\right) \right] \quad (2.18)$$

Original formula by De Boor uses the knots  $\{t_k\}$  but our definition of splines with even order is different.

Regular B-splines have a lot of interesting properties :

- the derivative of  $\beta_d$  is in the vectorial space generated by the shifted B-splines of order  $d-1$  :

$$\beta_d'(x) = \beta_{d-1}\left(x + \frac{1}{2}\right) - \beta_{d-1}\left(x - \frac{1}{2}\right) \quad (2.19)$$

- the scalar-product of two B-splines is easily computable :

$$\langle \beta_d(\cdot), \beta_{d'}(\cdot - \alpha) \rangle_{L^2(\mathbb{R})} = \beta_{d+d'+1}(\alpha) \quad (2.20)$$

- shifted B-splines  $\{\beta_d(\cdot - k)\}_{k \in \mathbb{Z}}$  form a *multi-resolution analysis* (see [65]) up to a translation of  $\frac{1}{2}$  for even orders. The inter-scale relation is given by :

$$\beta_d(x) = \frac{1}{2^d} \sum_{k=0}^{d+1} \binom{d+1}{k} \beta_d\left(2x - \frac{2k-d-1}{2}\right)$$

The signals we deal with are two dimensional (images), and are combination of translated functions  $\beta_d(\cdot - k)$  with different boundary conditions. We define the classical, periodic and symmetric-periodic approximation spaces for images with height  $m$  and width  $n$  :

$$S^d(n, m) = \left\{ \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \beta_{d,2}(\cdot - (k, l)) \text{ s.t. } (a_{k,l}) \in \mathbb{R}^{m \times n} \right\} \quad (2.21)$$

$$\begin{aligned} S_{per}^d(n, m) &= \left\{ \sum_{(p_1, p_2) \in \mathbb{Z}^2} f(\cdot + (p_1 n, p_2 m)) \text{ s.t. } f \in S^d(n, m) \right\} \\ &= \left\{ \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \tilde{\beta}_{d,2}(\cdot - (k, l)) \text{ s.t. } (a_{k,l}) \in \mathbb{R}^{m \times n} \right\} \end{aligned} \quad (2.22)$$

$$S_{sym}^d(n, m) = \left\{ f \in S_{per}^d(2n, 2m) \text{ s.t. } \begin{cases} f(-x, y) = f(x-1, y) \\ f(x, -y) = f(x, y-1) \end{cases} \forall x, y \right\} \quad (2.23)$$

where  $\tilde{\beta}_{d,2}$  is the periodized version of function  $\beta_{d,2} = \beta_d \otimes \beta_d$  with periods  $(n, 0)$  and  $(0, m)$ .  $S_{sym}^d(n, m)$  also admits a basis formed by periodization and then symmetrization relatively to axes  $\{x = -\frac{1}{2}\}$  and  $\{y = -\frac{1}{2}\}$ . This last spline space is related to symmetric boundary conditions in the direct spline transform as detailed in [90]. Interpolating B-splines, also called cardinal B-splines or fundamental splines, are basis functions denoted by  $(\eta_{d,2})_{k,l}$  such that for all  $f \in S^d(n, m)$  one has

$$f(\cdot) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} f(k, l) (\eta_{d,2})_{k,l}(\cdot)$$

Such a basis always exists in  $S^d(n, m)$ ,  $S_{per}^d(n, m)$  and  $S_{sym}^d(n, m)$  but the expression of the  $(\eta_{d,2})_{k,l}$  functions depends on the considered space. We define the spline transform and inverse spline transform operators (here on  $S^d(n, m)$ ) :

$$\mathcal{T}_d : f \rightarrow (a_{k,l}) \in \mathbb{R}^{n \times m} \text{ s.t. } f(\cdot) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \beta_{d,2}(\cdot - (k, l)) \quad (2.24)$$

$$\mathcal{T}_d^{-1} : (a_{k,l}) \in \mathbb{R}^{n \times m} \rightarrow \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \beta_{d,2}(\cdot - (k, l)) \quad (2.25)$$

We are generally interested in recovering a function in  $S^d(n, m)$  on the regular grid  $\Lambda_{reg}$ . Since a function in  $S^d(n, m)$  is uniquely determined by its values on  $\Lambda_{reg}$  we consider the spline sampling matrix  $M_d$  related to  $\mathcal{T}_d^{-1}$

$$(M_d)_{(k,l),(q,p)} = \beta_{d,2}(k - q, l - p) \quad (2.26)$$

We note that for the periodic space  $S_{per}^d(n, m)$  the spline sampling matrix is circulant and its inverse is given by inversion of its singular values in the Fourier basis of  $[0, n - 1] \times [0, m - 1]$ . In the other cases ( $S^d(n, m)$  and  $S_{sym}^d(n, m)$ ), a matrix inversion is necessary (by LU factorization for example) and a fast implementation is presented in [90].

The interpolating splines  $(\eta_{d,2})_{k,l}$  are given by

$$(\eta_{d,2})_{(k,l)}(\cdot) = \sum_{q=0}^{n-1} \sum_{p=0}^{m-1} (M_d^{-1})_{(k,l),(q,p)} \beta_{d,2}(\cdot - (q, p)) \quad (2.27)$$

In the periodic space  $S_{per}^d(n, m)$  the basis of interpolating splines is shift-invariant since  $(M_d)^{-1}$  is also circulant, but it is generally not the case in other spaces due to boundary conditions. The use of splines in signal approximation only makes sense if the sampling step (here 1) is adapted to the space of possible signals. As shown by Unser et al. [1], the cardinal spline  $\eta_{d,2}(\cdot)$  in the infinite setting converges to the cardinal sine function  $(\text{sinc}(\frac{\cdot}{\pi}))$  as  $d$  tends to infinity and thus provides a good approximation of bandlimited functions in  $BF(\pi)$ .

**Theorem 6 ([1] Aldroubi, Unser, Eden 1991)** *Let  $\eta_d$  be the cardinal spline on  $\mathbb{R}$  with order  $d \geq 0$ . The Fourier transform of  $\eta_d$  converges in  $L^p(\mathbb{R})$  to the indicator function of  $[-\pi, \pi]$  as  $d$  tends to infinity for all  $p \in [1, +\infty[$ .*

As a consequence, the sequence  $(f_d = \sum_{k \in \mathbb{Z}} c_k \eta_d(\cdot - k))_{d \geq 0}$  converges uniformly to the bandlimited function  $f \in BF(\pi)$  such that  $f(k) = c_k$  for all  $c \in \ell^2(\mathbb{Z})$  as  $d$  tends to infinity. In the finite setting  $S_{per}^d(n, m)$ , the cardinal spline function  $\eta_d$  converges to the Dirichlet kernel and the spline approximation converges to the trigonometric polynomial interpolation of the coefficients  $(c_k)$ .

## 2.2 Optimization in Hilbert spaces

Optimization is the field of mathematics which deals with the resolution of minimization problems

$$\text{find } x^* = \underset{x \in E \subset \mathcal{H}}{\operatorname{argmin}} f(x) \quad \text{or} \quad \text{compute } m = \min_{x \in E \subset \mathcal{H}} f(x) \quad (2.28)$$

All the properties and theorems in this section concern a real Hilbert space  $\mathcal{H}$ . When  $\mathcal{H}$  is finite-dimensional ( $\mathcal{H} = \mathbb{R}^d, d \geq 1$ ) we consider the canonical scalar product and norm

$$\langle x, y \rangle = \sum_{j=1}^d x_j y_j \quad , \quad \|x\| = \left( \sum_{j=1}^d x_j^2 \right)^{1/2} \quad x, y \in \mathbb{R}^d$$

We deal with cost/objective functions  $f : \mathcal{H} \rightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ . Conventions for combination of infinite values and/or scalars are described in [76] p.24 and the only forbidden case is " $\infty - \infty$ ".

### 2.2.1 Proper functions, Convexity, continuity, semi-continuity, inf-sup stability

Since we are interested in minimizing a cost function  $f$ , the values  $-\infty$  and  $+\infty$  have a special role. Nice functions for this problem are the ones which cannot take the  $-\infty$  value (in practice a cost or energy function always admits a finite minimum) and are not equal to  $+\infty$  everywhere (it would be useless to minimize such a function). These functions from  $\mathcal{H}$  to  $\mathbb{R} \cup \{+\infty\}$  are called *proper* function. The value  $+\infty$  plays a key role because the set of points where this value is reached forms an impossible set for the minimization problem. The set of possible minimizers, where  $f$  is not  $+\infty$  is called the domain of  $f$  and denoted by  $\operatorname{dom}(f)$ .

The main difficulty in finding the minimum of  $f$  with an algorithm is that numerical implementations generally use a local approach and can get stuck near a non-optimal local minimizer. A very important property to avoid this situation is the *convexity* property

**Definition 8 (Convex functions)** *A function  $f$  is said convex on  $\mathcal{H}$  if for all  $x, y \in \mathcal{H}$  and for all  $\theta \in ]0, 1[$ ,*

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$$

In this case two different local minima  $f(x^*)$  and  $f(y^*)$  such that  $f(x^*) < f(y^*)$  would verify for  $\theta \in ]0, 1[$ :

$$\begin{aligned} f(y^* + \theta(x^* - y^*)) &\leq \theta f(x^*) + (1 - \theta)f(y^*) \\ &< f(y^*) \end{aligned}$$

in contradiction with the local minimum property of  $y^*$ . Convexity of functions is closely related to convexity of sets :

**Definition 9 (Convex sets)** *A subset  $E \subset \mathcal{H}$  is said convex if for all  $x, y \in \mathcal{H}$  and for all  $\theta \in [0, 1]$*

$$\theta x + (1 - \theta)y \in E$$

Indeed, an equivalent definition for a convex function is a function with convex epigraph

$$\operatorname{epi}(f) = \{(x, \mu) \in \mathcal{H} \times \mathbb{R} \text{ s.t. } f(x) \leq \mu\}$$

Proper functions then correspond to functions with non-empty epigraph ( $f$  non identically  $+\infty$ ) which contain no vertical line (no  $x$  verifies  $f(x) \leq \mu$  for all  $\mu \in \mathbb{R}$  which is equivalent to  $f(x) \neq -\infty$ )

for all  $x \in H$ ). In this setting it is possible to extend any convex function  $f$  defined on a convex subset  $E$  of  $\mathcal{H}$  to the entire space by setting  $\tilde{f}(x) = f(x)$  on  $E$  and  $\tilde{f}(x) = +\infty$  on  $H \setminus E$  thanks to the convention

$$\forall \theta \in [0, 1], \forall C \in \mathbb{R} \cup \{+\infty\} \quad \theta C + (1 - \theta) \times (+\infty) = +\infty$$

The theory of convex analysis also treats the infinite dimensional case, which lead to the consideration of discontinuous convex functions. In the finite dimension case, finite convex functions behave locally like continuous functions.

**Proposition 4 (Rockafellar [76] p. 83)** *A function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is continuous if it is convex.*

Finiteness is important of course, consider for example the function  $\mathbb{I}$  defined on  $\mathbb{R}$  by

$$\mathbb{I}(x) = \begin{cases} 0 & \text{if } x \in [-1, 1] \\ +\infty & \text{else} \end{cases}$$

$\mathbb{I}$  is clearly not continuous at  $x = \pm 1$ . If we replace the interval  $[-1, 1]$  with any subset  $E$  of  $\mathcal{H}$  we call the function  $\mathbb{I}$  the *indicator* function of  $E$ , denoted by  $\mathbb{I}_E$ . Finiteness on  $\mathbb{R}^d$  and convexity imply continuity but a convex function is also continuous on any convex relatively open subset of its domain ([76] p. 82). The more general notion of lower semi-continuity is also very important in optimization.

**Definition 10 (Lower semi-continuity)** *A function  $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$  is said lower semi-continuous if its level sets  $s_\mu = \{x \in \mathcal{H} \text{ s.t. } f(x) \leq \mu\}$  are closed for all  $\mu \in \mathbb{R}$ .*

Full continuity is equivalent to lower and upper semi-continuity, where upper semi-continuity corresponds to the closedness of the sets  $s^\mu = \{x \in \mathcal{H} \text{ s.t. } f(x) \geq \mu\}$ . Lower semi-continuous functions have the following properties : if  $f$  is lower semi-continuous then for any sequence  $\{x_k\}$  converging to  $x^*$  with  $f(x^*) \neq +\infty$  we have

$$f(x^*) \leq \underline{\lim}_{k \rightarrow \infty} f(x_k)$$

such that if  $(f(x^k))_{k \geq 0}$  is a minimizing sequence of  $f$  and  $(x_k)_{k \geq 0}$  converges to  $x^*$  then

$$f(x^*) = \underline{\lim}_{k \rightarrow \infty} f(x_k) = \min_{x \in \mathcal{H}} f(x)$$

We also have that the supremum of a family of lower semi-continuous functions  $\{f_i\}_{i \in I}$  defined by  $S(x) = \sup_{i \in I} f_i(x)$  is a lower semi-continuous function and as well the supremum of a family of convex functions is a convex function. Lower semi-continuity and convexity are also preserved by the infimum  $I(x) = \inf_{i=1..p} f_i(x)$  over a finite family of functions. The indicator function above is an example of convex, lower semi-continuous function since  $s_\mu = [-1, 1]$  for  $\mu \geq 0$  and  $s_\mu = \emptyset$  else. An indicator function  $\mathbb{I}$  is convex if and only if the set  $E$  is convex and lower semi-continuous if and only if  $E$  is closed in  $\mathcal{H}$ . Proper, convex, semi-continuous functions on  $\mathcal{H}$  form an important class in Optimization theory denoted by  $\Gamma_0(\mathcal{H})$  ([29] p. 132). Note that finite convex functions are always lower semi-continuous, and that lower semi-continuous functions reach their minimum on any compact set of  $\mathcal{H}$ . If  $f \in \Gamma_0(\mathcal{H})$  is inf-compact ( $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ ) then it necessarily reaches its minimum.

### 2.2.2 Properties and characterization of convex functions

Convex functions in finite dimension are locally Lipschitz on the relative interior of their domain.

**Proposition 5 ([76] p. 86)** *Let  $f$  be a proper, convex function then  $f$  is Lipschitz on any relatively closed, bounded subset  $E$  of  $\text{ri}(\text{dom}(f))$  :*

$$\exists \alpha > 0 \quad \text{s.t.} \quad |f(x) - f(y)| \leq \alpha \|x - y\| \quad \text{for all } x, y \in E$$

where  $\text{ri}(\text{dom}(f))$  is the *relative interior* of  $\text{dom}(f)$  (see [76] p. 44). Consider the function  $g : x \rightarrow 1 - \sqrt{x(1-x)}$  on  $[-1, 1]$ , it is convex on  $[-1, 1]$  and can be extended to a convex function on  $\mathbb{R}$  but it is not Lipschitz on  $[-1, 1]$ , not even on  $] - 1, 1[$  but it is Lipschitz on any interval of the form  $[\alpha, \beta]$  with  $0 < \alpha \leq \beta < 1$ . In the case of a differentiable function  $f$  on an open subset  $E$ , convexity is equivalent to an inequality on the differences of  $f$  [10].

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle \quad \text{for all } x, y \in E \quad (2.29)$$

and in the case of a twice differentiable function, convexity is equivalent to the semi-positivity of the Hessian matrix  $H(x)$  on  $E$  with  $H_{j,k}(\cdot) = \frac{\partial^2 f}{\partial x_j \partial x_k}(\cdot)$ .

$$\langle V, H(x)V \rangle \geq 0 \quad \text{for all } V \in \mathcal{H} \text{ and for all } x \in E \quad (2.30)$$

A very close notion, *strict convexity*, concerns functions  $f$  such that for all  $x, y \in \text{dom}(f)$  and for all  $\theta \in ]0, 1[$

$$f(\theta x + (1 - \theta)y) < \theta f(x) + (1 - \theta)f(y)$$

and similar characterizations of convex differentiable or twice differentiable functions exist with a strict inequality in (2.29) and a positive definite Hessian in (2.30). Strict convexity assures uniqueness of the minimizer  $x^*$ . Convex functions on  $\mathbb{R}^d$  are not necessarily differentiable on their whole domain, but the set where a convex function is differentiable has good properties.

**Theorem 7 (Rockafellar [76] p. 246)** *Let  $f$  be a convex, proper function on  $\mathbb{R}^d$  and  $D$  the subset of  $\text{ri}(\text{dom}(f))$  where  $f$  is differentiable. Then  $D$  is a dense subset of  $\text{ri}(\text{dom}(f))$  and its complement in  $\text{ri}(\text{dom}(f))$  has measure zero. Furthermore the gradient  $\nabla f : D \rightarrow \mathbb{R}^d$  is continuous on  $D$ .*

We see here that convex functions in finite dimension are nice objects. When  $f$  is convex differentiable on  $\mathcal{H}$  the characterization of a minimizer  $x^*$  is quite easy

$$x = \underset{x \in \mathcal{H}}{\text{argmin}} f(x) \quad \Leftrightarrow \quad \nabla f(x) = 0$$

but a minimum can be reached at a nondifferentiable point of  $f$ , this is why we introduce the notion of subdifferential.

### 2.2.3 Subdifferential and Legendre transform

Let  $f$  in  $\Gamma_0(\mathcal{H})$ , at every point  $x$  it is possible to determine a set of directions which decrease  $f$  locally like  $-\nabla f(x)$  in the differentiable case.

**Definition 11 (Subdifferential)** *The subdifferential of  $f$ , denoted by  $\partial f$  is a set-valued operator from  $\mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$  defined by*

$$\partial f(x) = \{T \in \mathcal{H} \text{ s.t. } f(y) \geq f(x) + \langle T, y - x \rangle \text{ for all } y \in \mathcal{H}\}$$

If  $x \notin \text{dom}(f)$  then  $\partial f(x) = \emptyset$ . A vector in  $\partial f(x)$  is called a *subgradient* of  $f$  at  $x$  and determines an affine function tangent to  $f$  at  $x$ . The subdifferential  $\partial f(x)$  is a closed convex set, it is nonempty at points of  $\text{ri}(\text{dom}(f))$  and bounded at points of  $\text{Int}(\text{dom}(f))$  ([76] p. 217). It can be linked to the notion of Gâteaux derivative, since a subdifferentiable function at  $x \in \text{ri}(\text{dom}(f))$  is Gâteaux differentiable and the directional derivative  $f'(x, y) = \lim_{\alpha \rightarrow 0, \alpha > 0} \frac{1}{\alpha} (f(x + \alpha y) - f(x))$  verifies

$$f'(x, y) = \sup_{T \in \partial f(x)} \langle T, y \rangle$$

The subdifferential gives a necessary and sufficient condition for problem (2.28) :

**Proposition 6** *Let  $f \in \Gamma_0(\mathcal{H})$  and  $x^* \in \mathcal{H}$ . Then*

$$x^* = \underset{x \in \mathcal{H}}{\text{argmin}} f(x) \quad \Leftrightarrow \quad 0 \in \partial f(x^*)$$

When  $\partial f(x)$  is reduced to a single point the function  $f$  is differentiable at  $x$  and  $\partial f(x) = \{\nabla f(x)\}$ . Subdifferentiation of functions works (generally) like the classic differentiation operator.

**Proposition 7 (Rockafellar [76] p. 225)** *Let  $f, g \in \Gamma_0(\mathcal{H})$ ,  $A$  a linear mapping from  $\mathbb{R}^p \rightarrow \mathbb{R}^d$ .*

$$\partial(f + g)(x) \supset \partial f(x) + \partial g(x) \quad \text{for all } x \in \text{dom}(f + g)$$

*with equality if  $\text{ri}(\text{dom}(f)) \cap \text{ri}(\text{dom}(g)) \neq \emptyset$ , and*

$$\partial(f \circ A)(x) \supset A^* \partial f(Ax) \quad \text{for all } x \in \text{dom}(f \circ A)$$

*with equality if  $R(A) \cap \text{ri}(\text{dom}(A)) \neq \emptyset$ .*

These conditions are always verified by the functions we deal with. A duality principle states that a function in  $\Gamma_0(\mathcal{H})$  is the supremum of all affine functions majorized by  $f$  and in particular the affine functions of the form  $x \rightarrow \langle x, x^* \rangle - \mu^*$  whenever  $\mu^* \geq \sup_{z \in \mathcal{H}} \langle z, x^* \rangle - f(z)$ . This function of  $x^*$  is called the *conjugate* or *Legendre transform* of  $f$  at  $x^*$  and it is a one-to-one mapping of  $\Gamma_0(\mathcal{H}) \rightarrow \Gamma_0(\mathcal{H})$ .

**Proposition 8 (One-to-one Legendre transform)** *Let  $f \in \Gamma_0(\mathcal{H})$ , there exists a unique function  $f^* \in \Gamma_0(\mathcal{H})$  such that*

$$f(x) = \sup_{x^* \in \mathcal{H}} \langle x^*, x \rangle - f^*(x^*) \quad \text{for all } x \in \mathcal{H}$$

*and the conjugate of  $f^*$  is  $f$  itself*

$$(f^*)^* = f$$

The relation between conjugacy and subdifferential is :

$$\partial f(x) = \{y \in \mathcal{H} \text{ s.t. } f(x) + f^*(y) = \langle x, y \rangle\}$$

such that

$$y \in \partial f(x) \quad \Leftrightarrow \quad x \in \partial f^*(y)$$

This provides an alternative to the problem in  $f$  :

$$0 \in \partial f(x) \quad \Leftrightarrow \quad x \in \partial f^*(0)$$

### 2.2.4 Lagrangian, duality gap, Fenchel-Moreau duality

Minimization problems often contain one or several constraints :

$$x^* = \operatorname{argmin}_{x \in \mathcal{H}} f(x) \quad \text{s.t.} \quad \begin{cases} g_i(x) = 0 & \text{for } 1 \leq i \leq N_e \\ h_j(x) \leq 0 & \text{for } 1 \leq j \leq N_i \end{cases} \quad (P_{E,I})$$

with  $N_e$  the number of equality constraints and  $N_i$  the number of inequality constraints. In the case when the  $g_i$  functions are affine functions and the  $h_j$  functions are finite functions in  $\Gamma_0(\mathcal{H})$ , we introduce the notion of *Lagrangian function*  $\mathcal{L} : \mathcal{H} \times \mathbb{R}^{N_e} \times \mathbb{R}^{N_i} \rightarrow \mathbb{R} \cup \{+\infty\}$  associated to  $(P_{E,I})$  such that

$$\sup_{(\lambda, \mu) \in \mathbb{R}^{N_e} \times \mathbb{R}^{N_i}} \mathcal{L}(x, \lambda, \mu) = \begin{cases} f(x) & \text{if } x \text{ satisfies } (P_{E,I}) \\ +\infty & \text{else} \end{cases}$$

A Lagrangian function transforms the minimization problem  $(P_{E,I})$  in  $x$  into a saddle point problem in  $(x, \lambda, \mu)$

$$x^* = \operatorname{argmin}_{x \in \mathcal{H}} \sup_{(\lambda, \mu)} \mathcal{L}(x, \lambda, \mu) \quad (2.31)$$

A classical example of Lagrangian function is

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{N_e} \lambda_i g_i(x) + \sum_{j=1}^{N_i} \mu_j h_j(x) \quad \lambda \in \mathbb{R}^{N_e}, \mu \in (\mathbb{R}_+)^{N_i}$$

The saddle point problem (2.31) is called the *primal problem* and inf and sup inversion gives the *dual problem*

$$\sup_{(\lambda, \mu)} \min_{x \in \mathcal{H}} \mathcal{L}(x, \lambda, \mu) \quad (2.32)$$

This provides an empirical method for solving  $(P_{E,I})$  : first minimize the Lagrangian in  $x$  and then explore the space of Lagrange multipliers  $(\lambda, \mu)$ . When there are only inequality constraints, this approach can be seen as a penalization method on  $f$  with penalization functions  $h_j$  : the higher the penalization parameter  $\mu_j$ , the lower the value of  $h_j(x^*)$ . The Karush-Kuhn-Tucker theorem ([76] p. 293) provides an existence result for the Lagrange parameters  $(\lambda, \mu)$  to problem (2.31) whenever the original problem  $(P_{E,I})$  is feasible, that is there exists  $x^* \in \mathcal{H}$  such that all the constraints hold. This duality principle can be applied to general saddle point problems providing that both primal and dual problems have the same value :

$$\left( \inf_{x \in X} \sup_{s \in S} F(x, s) \right) - \left( \sup_{s \in S} \inf_{x \in X} F(x, s) \right) = 0$$

This quantity is always positive and called the *duality gap*. It is not always zero and attention should be paid to the inf-sup exchange procedure. The inf-sup exchange in the Fenchel-Rockafellar duality formula is a case of zero duality gap.

**Theorem 8 (Fenchel-Rockafellar duality principle [30])** *Let  $f \in \Gamma_0(\mathbb{R}^d)$ ,  $g \in \Gamma_0(\mathbb{R}^p)$  and  $A$  a linear operator  $\mathbb{R}^d \rightarrow \mathbb{R}^p$  such that  $0 \in \operatorname{Int}(\operatorname{dom}(g) - A \operatorname{dom}(f))$ . Then*

$$\inf_{x \in \mathbb{R}^d} f(x) + g(Ax) = -\min_{s \in \mathbb{R}^d} f^*(-A^*s) + g^*(s)$$

### 2.2.5 The proximal operator, examples

We use intensively in this thesis minimization problems involving the sum of two convex functions

$$\min_{x \in \mathcal{H}} f(x) + g(x)$$

with  $f$  convex nondifferentiable and  $g$  convex differentiable. The subdifferential equation of this problem can be solved by a semi-implicit scheme [28]

$$\begin{cases} y^{n+1} &= x^n - \rho \nabla g(x^n) \\ y^{n+1} &\in (Id + \rho \partial f)(x^{n+1}) \end{cases}$$

The second step of this scheme always has a unique solution  $x^{n+1}$  and the operator defined by this inclusion is called the *proximal operator* of  $(\rho f)$ .

**Proposition 9 (Proximal operator)** *Let  $f \in \Gamma_0(H)$  and  $\rho > 0$ . For all  $x \in \mathcal{H}$ , there exists a unique vector  $y$  such that  $x \in y + \rho \partial f(y)$ , called the proximal point of  $f$  at  $x$  with parameter  $\rho$ , denoted by  $\text{prox}_{\rho f}(x)$  and characterized by the variational problem*

$$y = \text{prox}_{\rho f}(x) \iff y = \underset{z \in \mathcal{H}}{\text{argmin}} f(z) + \frac{1}{2\rho} \|x - z\|^2$$

We present in table 2.1 examples of proximal operators. The proximal operator has many properties, among which the decomposition formula

**Proposition 10** *For all  $f \in \Gamma_0(\mathcal{H})$  and  $x \in \mathcal{H}$ ,*

$$x = \text{prox}_f(x) + \text{prox}_{f^*}(x)$$

It is firmly nonexpansive and thus continuous

$$\|\text{prox}_f(x) - \text{prox}_f(y)\|^2 \leq \langle \text{prox}_f(x) - \text{prox}_f(y), x - y \rangle \quad \text{for all } x, y \in \mathcal{H}$$

Functions of the type  $\sup_{s \in S} \langle Ax, s \rangle$  (with  $A$  a linear mapping  $\mathbb{R}^d \rightarrow \mathbb{R}^p$ ) are special functions with the *positive homogeneous* property

**Definition 12 (Positive homogeneous functions)** *A function  $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$  is positive homogeneous if for all  $x \in \mathcal{H}$  and  $\alpha > 0$*

$$f(\alpha x) = \alpha f(x)$$

These functions rewrite as the conjugate of an indicator function  $\mathbb{I}_K$  where  $K$  is a convex closed subset of  $\mathcal{H}$ , namely the zero level of the conjugate :

**Proposition 11 (Rockafellar [76] p. 114)** *Let  $f \in \Gamma_0(\mathcal{H})$  a positive homogeneous function,  $K = s_0(f^*)$  the zero levelset of  $f^*$ , then*

$$f = (\mathbb{I}_K)^*$$

*and  $K$  is bounded if and only if  $f$  is finite.*

Function	Proximal Operator
$f(x) = \mathbb{I}_E(x)$ with $E$ a closed convex subset of $\mathcal{H}$	$\text{prox}_{\rho f}(y) = \text{proj}_E(y)$
$g(x) = \sup_{s \in K} \langle Ax, s \rangle$	$\text{prox}_{\rho g}(y) = y - \text{proj}_{\rho A^* K}(y)$
$h(x) = \frac{1}{2} \ Ax - z\ ^2$	$\text{prox}_{\rho h}(y) = (Id + \rho A^* A)^{-1} (y + \rho A^* z)$
$k(x) = \ x\ _1 = \sum_{j=1}^d  x_j $ on $\mathbb{R}^d$	$(\text{prox}_{\rho k}(y))_j = 0$ if $ y_j  \leq \rho$ $y_j - \frac{\rho}{ y_j } y_j$ else
$l(x) = \sum_{j=1}^d  x_j $ on $\mathbb{C}^d$	$z = \underset{x \in \mathbb{C}^d}{\text{argmin}} \quad l(x) + \frac{1}{2\rho} \sum_{j=1}^d  x_j - y_j ^2$ then $z_j = 0$ if $ y_j  \leq \rho$ and $y_j - \frac{\rho}{ y_j } y_j$ else

Table 2.1: Proximal operators examples.

The  $\ell^1$ -norm belongs to this class and we note that the set  $S$  in this case is  $[-1, 1]^d$ . Another classical example is the discrete total variation  $TV(f) = \|\nabla f\|_1$ , it is the conjugate function of  $\mathbb{I}_{\nabla^* S}$  where  $S = \left\{ s \in (\mathbb{R}^2)^d \text{ s.t. } \|s_j\| \leq 1 \text{ for all } j \in \{1..d\} \right\}$  such that

$$\text{prox}_{\rho TV(\cdot)}(y) = y - \text{prox}_{\mathbb{I}_{\nabla^* S}}(y) = y - \text{proj}_{\rho \nabla^* S}(y)$$

The existence and uniqueness of the orthogonal projection in Hilbert space are always verified thanks to the projection theorem.

**Theorem 9 (Projection theorem)** *Let  $\mathcal{H}$  be a Hilbert space with scalar product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  and norm  $\|\cdot\|_{\mathcal{H}}$ . Let  $E$  be a non-empty, convex, bounded subset of  $\mathcal{H}$  then for all  $z \in \mathcal{H}$  the following problem*

$$\min_{x \in E} \|x - z\|_{\mathcal{H}}$$

*admits a unique solution denoted by  $x^* = \text{proj}_E(z)$  and characterized by*

$$\text{Re}(\langle z - x^*, x - x^* \rangle_{\mathcal{H}}) \leq 0 \quad \text{for all } x \in E$$

## 2.2.6 Minimization algorithms and convergence results

### 2.2.6.1 The Conjugate Gradient algorithm

Quadratic functions on  $\mathbb{R}^d$  of the form  $F(x) = \frac{1}{2} \langle Ax, x \rangle + \langle y, x \rangle$  are frequently used in optimization algorithms like Quasi-Newton methods or quadratic regularization. A function of this form

with  $A \in \mathcal{M}_d(\mathbb{R})$  is strictly convex if and only if  $A + A^t$  is a positive definite matrix. This gives a variational foundation to the solving of linear systems  $Ax = y$  with  $A$  positive definite, very helpful for a line search descent. This line search, together with an orthogonalization of residues results in the very famous Conjugate Gradient method detailed below (algorithm 2.2).

Let  $y \in \mathcal{H}$ ,  $A \in \mathcal{M}_d(\mathbb{R})$  a positive definite matrix. Let  $(x_n), (p_n), (q_n) \in \mathcal{H}$ ,  $(\alpha_n), (\beta_n) \in \mathbb{R}$  be finite sequences defined by :

- $x_0 \in \mathcal{H}$ ,  $r_0 = y - Ax_0$ ,  $p_0 = r_0$ ,  $q_0 = Ap_0$
- $\alpha_0 = \frac{\|r_0\|^2}{\langle p_0, q_0 \rangle}$ ,  $x_1 = x_0 + \alpha_0 p_0$ ,  $r_1 = r_0 - \alpha_0 q_0$
- For all  $n \geq 1$  such that  $\|r_n\| > \epsilon$  :
  - $\beta_n = \frac{\|r_n\|^2}{\|r_{n-1}\|^2}$
  - $p_n = r_n + \beta_n p_{n-1}$  (conjugated descent directions)
  - $q_n = Ap_n$
  - $\alpha_n = \frac{\|r_n\|^2}{\langle p_n, q_n \rangle}$  (line search step)
  - $x_{n+1} = x_n + \alpha_n p_n$  (optimal gradient descent)
  - $r_{n+1} = r_n - \alpha_n q_n$  (residue update)

Figure 2.2: The (classical) Conjugate Gradient algorithm for symmetric positive definite linear systems.

This algorithm guarantees that  $x_n$  minimizes  $F$  over the affine subspace  $x_0 + K_n$  with  $K_n = \text{vect}(p_0, \dots, p_n)$  the Krylov space at iteration  $n$ . Residues  $r_n$  and descent directions  $p_n$  are conjugate relatively to  $A$ .

$$\langle r_n, Ar_k \rangle = 0 \quad \text{for all } 0 \leq k < n$$

It converges (theoretically) to a solution of  $Ax = y$  in at most  $d$  iterations (there exists  $d_0 \leq d$  such that  $r_{d_0} = 0$ ) but it is also used as an iterative algorithm for sparse linear systems with good conditioning.

**Theorem 10** *Let  $A$  be a positive definite matrix on  $\mathbb{R}^d$ ,  $y \in \mathbb{R}^d$  and  $x^*$  the solution of  $Ax = y$ . Let  $\kappa$  be the condition number of  $A$  and  $(x_n)$  a finite sequence defined by the previous CG algorithm 2.2. For all  $n \geq 0$  such that  $x_n$  is defined*

$$\|x_n - x^*\| \leq 2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^n \|x_0 - x^*\|$$

### 2.2.6.2 Nonsmooth optimization : Forward-Backward, FISTA, TWIST, Nesterov

In the setting of nonsmooth+smooth minimization, the classical semi-implicit scheme also called *Forward-Backward* amounts to compute iteratively an explicit descent step  $y^{n+1} = x^n - \rho \nabla g(x^n)$  and an implicit inclusion

$$y^{n+1} \in (Id + \rho \partial f(x^{n+1})) \Leftrightarrow x^{n+1} = \underset{z \in \mathcal{H}}{\text{argmin}} \quad \rho f(z) + \frac{1}{2} \|y^{n+1} - z\|^2$$

We recognize here the proximal of  $f$ . When  $f = \lambda \|x\|_1$  this algorithm is called ISTA (Iterative Soft-Thresholding Algorithm) but this denomination is also used for the general Forward-Backward algorithm. Convergence of the Forward-Backward algorithm (2.3) was proved with optimal step by Combettes [28].

- Let  $x_0 \in \mathcal{H}$ ,  $L(g)$  the Lipschitz constant of  $\nabla g$  and  $0 < \rho < \frac{2}{L(g)}$
- For all  $n \geq 0$  :
  - $x_{n+1/2} = x_n - \rho \nabla g(x_n)$
  - $x_{n+1} = \text{prox}_{\rho f}(x_{n+1/2})$

Figure 2.3: Scheme of the Forward-Backward algorithm with nonsmooth  $f$  and differentiable  $g$ 

**Theorem 11** ([30] **Combettes, Wajs 2009**) *Let  $f, g \in \Gamma_0(\mathbb{R}^d)$  with  $g$  differentiable on  $\mathbb{R}^d$ ,  $\nabla g$  Lipschitz on  $\mathbb{R}^d$ ,  $L(g) > 0$  its Lipschitz constant and  $0 < \rho < \frac{2}{L(g)}$ . The sequence  $(x_k)_{k \geq 0}$  defined by and*

$$x_0 \in \mathcal{H} \quad x_{n+1} = \text{prox}_{\rho f}(x_n - \rho \nabla g(x_n)) \quad \text{for all } n \in \mathbb{N}$$

*converges weakly and thus strongly to a minimizer of  $f + g$ .*

The convergence speed of ISTA was studied by Beck et al. [8] and proved a  $O(\frac{1}{N})$  (sublinear) with  $N$  the number of iterations.

**Theorem 12** ([8] **Beck, Teboulle 2009**) *Let  $f, g$  be convex functions on  $\mathbb{R}^d \rightarrow \mathbb{R}$ , with  $g$  differentiable and  $\nabla g$  Lipschitzian with constant  $L(g)$ . Let  $X^*$  be the set of minimizers of  $F = f + g$ , the sequence  $(x_n)_{n \in \mathbb{N}}$  defined in theorem 11 with  $\rho \leq \frac{1}{L(g)}$  verifies for all  $n \geq 1$  and  $x^* \in X^*$*

$$F(x_n) - F(x^*) \leq \frac{L(g) \|x_0 - x^*\|^2}{2n}$$

Bioucas-Dias et al. proposed the TWIST algorithm (TWo step Iterative Soft-Thresholding) as an accelerated scheme for the ISTA algorithm, and Beck et al. proposed the FISTA algorithm (Fast ISTA) which has a  $O(\frac{1}{n^2})$  convergence rate. These accelerated algorithms are not necessarily monotone, which is a big disadvantage in numerical optimization. The TWIST algorithm and its monotone version are described in [12], here we only present the FISTA and monotone FISTA algorithms we used in this thesis.

- Let  $x_0 \in \mathcal{H}$ ,  $y_1 = x_0$ ,  $t_1 = 1$ ,  $L(g)$  the Lipschitz constant of  $\nabla g$  and  $0 < \rho \leq \frac{1}{L(g)}$
- For all  $n \geq 1$  :
  - $x_n = \text{prox}_{\rho f}(y_n - \rho \nabla g(y_n))$
  - $t_{n+1} = \frac{1 + \sqrt{1 + 4(t_n)^2}}{2}$
  - $y_{n+1} = x_n + \left(\frac{t_n - 1}{t_{n+1}}\right) (x_n - x_{n-1})$

Figure 2.4: Scheme of the FISTA algorithm for the nonsmooth+smooth minimization of  $f + g$

- Let  $x_0 \in \mathcal{H}$ ,  $y_1 = x_0$ ,  $t_1 = 1$ ,  $L(g)$  the Lipschitz constant of  $\nabla g$  and  $0 < \rho \leq \frac{1}{L(g)}$
- For all  $n \geq 1$  :
  - $z_n = \text{prox}_{\rho f}(y_n - \rho \nabla g(y_n))$
  - $t_{n+1} = \frac{1 + \sqrt{1 + 4(t_n)^2}}{2}$
  - $x_{n+1} = \text{argmin} \{F(x) \text{ s.t. } x \in \{x_n, z_n\}\}$
  - $y_{n+1} = x_n + \left(\frac{t_n}{t_{n+1}}\right) (z_n - x_n) + \left(\frac{t_n - 1}{t_{n+1}}\right) (x_n - x_{n-1})$

Figure 2.5: Monotone version of FISTA

**Theorem 13** ([7] Beck and Teboulle 2009) *Let  $f, g$  be convex functions  $\mathbb{R}^d \rightarrow \mathbb{R}$  with  $g$  differentiable and  $\nabla g$  Lipschitz.*

*Let  $(x_n)_{n \in \mathbb{N}}$  the sequence generated by either FISTA (algorithm 2.4) or mFISTA (algorithm 2.5). For all  $x^* \in X^*$  the set of minimizer of  $F$  and for all  $n \geq 1$*

$$F(x_n) - F(x^*) \leq \frac{2L(f) \|x_0 - x^*\|^2}{(n+1)^2}$$



# Chapter 3

## Estimation of the microvibrations

... of a satellite in a couple of push-broom images

### 3.1 Exploitation of the disparity map

#### 3.1.1 Introduction

Estimation of the microvibrations of a satellite is a very important point in sub-pixel disparity computation. Aliasing and noise reduction in modern Earth imaging devices have made possible the reconstruction of 3D scenes from small B/H image couples (B is the baseline or instrument displacement, H is its altitude) with high precision [35], and this small B/H acquisition mode will be used by the PLEIADES satellites. Depth estimation in an image couple is a large field of computer vision, and the most famous method is normalized cross-correlation (NCC) minimization along epipolar lines<sup>1</sup> (or at a neighborhood), or the simple sum of squared differences (SSD) :

$$d_{u_1}(x) = \min_y \iint_{\mathbb{R}^2} \phi(y) |u_1(x) - u_2(x+y)|^2 dy \quad (3.1)$$

where  $u_1, u_2$  are the two images and  $\phi$  is a two-dimensional window function, generally a truncated Gaussian function with small standard deviation. These methods require rectified images, which consist in a continuous scene projected on the same plane but from different optical centres. This can be achieved by resampling the second image  $u_2$  on a grid determined by the calibration parameters of both cameras. If the parameters or some of them are unknown (for example the exact location and angular state of the cameras), then one must address the bundle adjustment problem or estimation of the transformation between the two images. Once the images are correctly stereo-rectified, the computation of the disparity in image  $u_1$  by the minimization (3.1) can be performed but suffers from several errors :

- adhesion or dilatation : when elevation varies quickly the image contents can change very fast (contraction) or even disappear (occlusion) and the correlation of elevated points (closer to the imaging device) can be associated to less elevated ones.
- stroboscopic effect : each part of a repetitive pattern is likely to match with any other part of the pattern.

---

<sup>1</sup>the epipolar line of a point A in image 1 is the projection in image 2 of the set of 3D points whose projection in image 1 is A

- wrong matches in under-detailed areas and/or noisy areas (noise has a bigger influence in the smooth parts of an images).

These phenomena are well described in [82], as well as the literature in stereo-vision. The algorithm proposed in [82] is now implemented in the CNES MARC2 software (Multiresolution Algorithm for Refined Correlation) used in this report for the computation of the disparity maps<sup>2</sup>. The reason why we focus on disparity is that there is no easy way to estimate the deformation in a single image without any a priori on the ground truth : roads are not exactly straight, the size and spacing of buildings are not always available and the placement of mires on the ground with appropriate spacing is too demanding and rarely possible. Automatic estimation of the sampling error necessarily involves two views of the same scene and the intuitive approach for this estimation is a comparison of the locations of features present in both images.

### 3.1.2 State of the art

Influence of microvibrations in Earth images is illustrated in figure 3.1 where we display the original images and the images with deformation. The microvibrations oscillate between in  $[-1, 1]$  and have a negligible influence to a human eye but modify the disparity map. The microvibrations add an oscillating component to the disparity map.

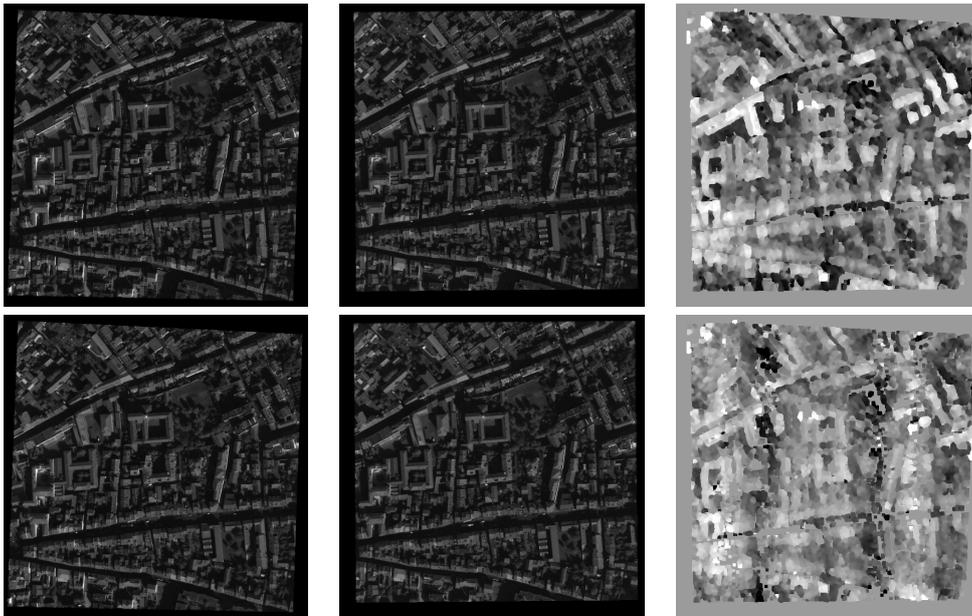


Figure 3.1: Without (top row) and with microvibrations (second row). The difference between images of the first and second row is invisible to the human eye although the disparity map is perturbed by low frequency, horizontal oscillations.

A simple method used to cancel this kind of oscillating perturbations is FFT selective filtering [63], this is a thresholding method in the Fourier domain. In this approach the removal of the perturbation consists in the application of a mask to the Fourier coefficients. This implies that

<sup>2</sup>We thank the CNES for providing this software.

some frequencies are completely removed and the amount of perturbation removed this way is not constant in all parts of the disparity map. The authors of [63] addressed here the more general denoising problem in which the noise has a small support in the Fourier domain and unknown amplitude. In our case it would be possible to estimate the perturbation in homogeneous areas as in the selective FFT filtering method, and then subtract this perturbation but :

- the selected area should have a size of several periods of the microvibration (which is not possible in urban landscapes).
- the true elevation of the selected area may not verify our homogeneous assumption, even after a vertical smoothing (if the elevation is affine for example)

In a work by Roques et al. [78] (2001), the authors estimate the three angular vibrations of a satellite with dense disparity maps using various methods. As in the FFT selective filtering method, the pitch is estimated from manually selected homogeneous areas. This estimation is completed by a Matching Pursuit algorithm (see section 3.2.1) in a redundant time-frequency dictionary.

The microvibration problem has also been treated at CNES by Breton et al. [17] and then Greslou et al. [45] for SPOT5 HRG images with several couples from different sensors (the SPOT5 HRG device is made of a panchromatic and 4 spectral sensors). In [45], the authors work with quasi-stationary and quasi-harmonic noisy signals, and modelize the microvibrations as a sum of sine functions depending on the time variable and shifted by a factor  $\tau_j$  where  $j$  is the index of the image couple :

$$\begin{cases} V_{1,j}(\cdot) &= \sum_{k=1}^n a_k \sin(w_k \cdot + \phi_k) \\ V_{2,j}(\cdot) &= V_1(t + \tau_j) \end{cases} \quad (3.2)$$

The different  $\tau_j$  parameters depend on the shape and placement of the sensors and lead to an overdetermined linear system solved by Least-Squares method.

### 3.1.3 Framework of this study

On the contrary of [45], we work here with a single couple of stereo-rectified images, with high resolution and quasi-harmonic microvibrations as modeled at CNES (see table 3.0(a)). We studied both stationary and non-stationary cases : the case of non-stationary microvibrations occurs with agile devices which alternatively point forward during acquisition of the first image and backward once the satellite has passed over the target. This repointing has an influence on the angular oscillations and microvibrations in the sampling grid of the two images can be decorrelated. In the stationary case (no repointing) we assume that the microvibrations  $V_1$  and  $V_2$  are correlated up to a time shift ( $V_2(t) \approx V_1(t + \tau)$ ).

Our experiments only concern the correction of pitch oscillations, in the case of a sensor line orthogonal to the flight direction as described in section 1.1.2. We tried to improve the approach in [78] when only two images of the same scene are available. The available data is a lacunary disparity map and we approximate the microvibrations with sparse signals in the Fourier domain. In a second part (section 3.4) we add an a priori on the elevation map in order to fix some incorrect detections and show the improvements in urban landscapes.

### 3.1.4 The MARC2 software package

MARC software has been developed at CNES and is continuously improved by new techniques from image processing and stereo-vision. The version we present here benefits from the recent works of Sabater et al. [81] [82] [83]. The software was used with input data consisting in two

stereo-rectified views from the same scene  $I_1$  and  $I_2$  suffering from microvibrations and produced several output disparity maps.

The package implements an automatic rejection of false matches thanks to an a contrario approach and uses only a very small number of input parameters. The disparity associated to a point in image  $I_1$  is a measure of its elevation obtained by subtracting its (supposed) location in  $I_2$  to its coordinates in  $I_1$ , it can be computed up to a very small precision  $\approx \frac{1}{20}$  pixel with high quality images such as PLEIADES images but only certain regions are assigned a disparity measure, typically the ones containing structures.

A first set of positive matches is determined using a Principal Component Analysis on classes of patches (neighbourhood of a point) with similar properties (mean and variance). Patches from a patch class are projected on a set of orthonormal components, namely the principal components of the class, this gives an empirical distribution for PCA coefficients. As remarked by many authors, this set of principal components is almost independent of the image and can be computed once for all.

The disparity is determined as the minimum of a correlation function  $d_{I_1}$  at point  $M_0 = (x_0, y_0)$  belonging to  $I_1$ , a distance between the patch centered on  $M_0$  and all other patches centered on  $M(x, y_0)$  on the same line as  $M_0$  in image  $I_2$ .

$$d_{I_1}(M_0, x) = \iint_V I_1(x_0 + s, y_0 + t) I_2(x + s, y_0 + t) \phi(s, t) ds dt \quad (3.3)$$

with  $V$  a neighbourhood of  $(0, 0)$  and  $\phi$  a window function.

The first elimination step of the algorithm is based on the a contrario approach [37] [82] and performed before the minimization of  $d_{I_1}(M_0, x)$ . This gives a first disparity map with pixel precision (disp1). The disparity is then refined with subpixel precision (disp\_sub) by minimizing 3.3 with bandlimited hypothesis on  $I_1$  and  $I_2$ . It is then made denser by the application of a median filter (disp\_med) and undergoes an anti-adhesion filtering process based on line segments detection (disp\_final).

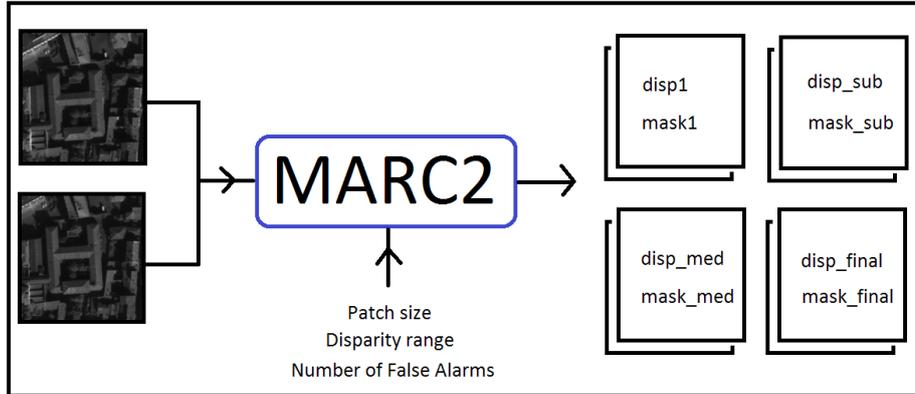


Figure 3.2: Principle of MARC2 software. The output disparity map (the masks correspond to binary images with value 1 where the disparity is valid).

The information available in these different disparity maps is thus one-dimensional since only patches on the same horizontal line are compared. In our case the  $y$  displacement information is not necessary as we consider horizontal microvibrations ( $\epsilon_y = 0$ ) but the algorithm can be applied

to stereo-rectified images with 2D microvibrations if necessary with the risk for increased rejections from the a contrario method.

### 3.1.5 Information in the disparity map

The MARC2 software produces several disparity maps with variable density, depending on the number of false alarms in the a contrario method, the precision of the disparity and the correction of adhesion. This density also depends on the contents of images (see figure 3.3) and we considered in our tests the final disparity map, which is even less dense than figure 3.3.

We denote by  $(d_k = d(z_k))_{1 \leq k \leq |\Lambda|}$  the vector of known disparities on the grid

$$\Lambda = \{z_k = (x_k, y_k)\}$$

where the mask given by MARC2 is equal to 1 (each  $d_k$  is a 2 dimension vector but here the vertical component is assumed to be zero).  $\Lambda$  is a subset of  $[0, m - 1] \times [0, n - 1]$  with integer coordinates (examples are displayed in the last column of figure 3.3) and the disparities  $d_k$  correspond to the measured displacement values of points in image  $I_1$ . This data can be expressed as a function of the true displacement  $h$  and the microvibrations in each image  $V_1$  and  $V_2$ . In the following, the continuous functions  $I_1$  and  $I_2 : [0, m[ \times [0, n[ \rightarrow \mathbb{R}$  correspond to the interpolation of the left and right images  $I_1$  and  $I_2$ . We also introduce two continuous signals  $U_1$  and  $U_2$ , the continuous scenes observed in the rectification plane from each optical center. Images  $I_1$  and  $I_2$  have the expression :

$$I_i(z) = U_i(z + V_i(z)) \quad i \in \{1, 2\} \quad (3.4)$$

We note  $(h_k = h(z_k))_{z_k \in \Lambda}$  the true displacements of the physical points in image 1, while the  $(d(z_k))_{z_k \in \Lambda}$  are the measured displacements :

$$\begin{cases} I_1(z_k) = I_2(z_k + d(z_k)) & \text{for all } 1 \leq k \leq |\Lambda| \\ U_1(z) = U_2(z + h(z)) & \text{locally in } z \end{cases} \quad (3.5)$$

This gives the expression of  $d_k$  :

$$\begin{aligned} I_1(z_k) &= U_1(z_k + V_1(z_k)) \\ &= U_2(z_k + V_1(z_k) + h(z_k + V_1(z_k))) \end{aligned}$$

On the other side we have

$$\begin{aligned} I_1(z_k) &= I_2(z_k + d_k) \\ &= U_2(z_k + d_k + V_2(z_k + d_k)) \end{aligned}$$

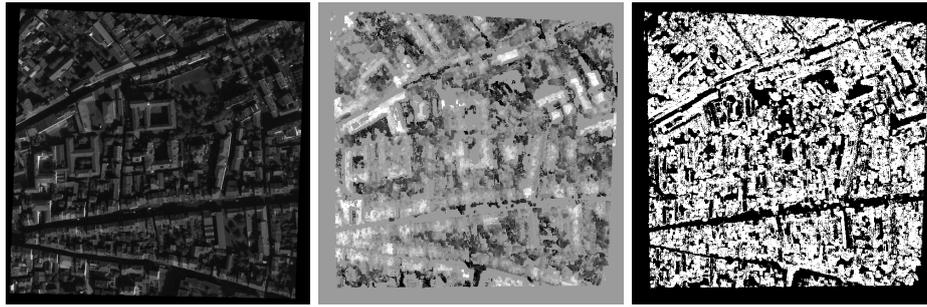
$$\boxed{V_1(z_k) + h(z_k + V_1(z_k)) = d_k + V_2(z_k + d_k)}$$

Finally, in order to facilitate the interpretation of the data, we consider the opposites of  $d_k$  and  $h_k$  such that objects close to the rectification plane have a larger disparity :

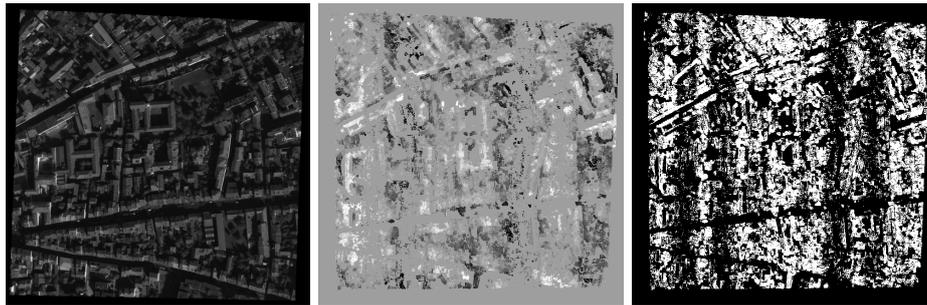
$$\begin{cases} D_k = -d_k \\ H_k = -h_k \end{cases}$$

$$\Rightarrow \boxed{D_k = H(z_k + V_1(z_k)) - V_1(z_k) + V_2(z_k - D_k)} \quad (3.6)$$

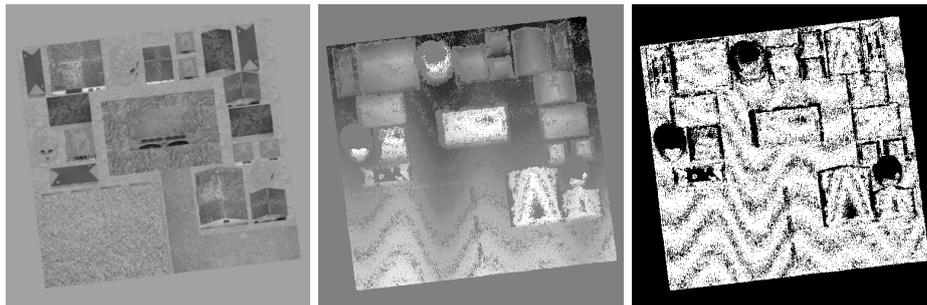
Equation (3.6) verified by the measured disparities shows the mixing between  $V_1$  and  $V_2$ , although one microvibration is sampled at integer locations and the other one is sampled on the irregular



(a) Disparity from simulated PLEIADES images (CNES copyrighted image) (middle) and corresponding mask (right) (no microvibrations).



(b) In the presence of microvibrations.



(c) Disparity of *Village* image from L. Moisan's images simulator (middle) and corresponding mask (right)

Figure 3.3: Examples of disparity maps from MARC2 algorithm (the one with the highest density possible here). Resolution and the presence of deep black areas have a great influence on the density of the disparity maps (first row). The map can be even less dense in the presence of microvibrations (second row)

grid  $\Lambda' = \{z_k - D_k\}_{1 \leq k \leq |\Lambda|}$ . The other point, as expected, is that the real disparity  $H$  is also sampled on an irregular grid  $\{z_k + V_1(z_k)\}_{1 \leq k \leq |\Lambda|}$  and can be corrected only if  $V_1$  is well estimated.

Our goal here is to retrieve  $H$  from  $D$  without any a priori on the observed scene, except that it does not look like a low frequency sine function. By this assumption we mean that the horizontal component of the elevation  $H$  should be homogeneously distributed in the Fourier domain.

## 3.2 Detection by $l^1$ regularization

### 3.2.1 A nonlinear solver for sparse signals

In the late 90's, researchers from the Signal Processing community addressed a special kind of problem : reconstruction of signals from only a few measurements. In some cases, the physical information could not be acquired at the Nyquist's rate given by Shannon's theorem and it was necessary to develop new algorithms. The first methods were computationally heavy and called *Greedy Algorithms* : Matching Pursuit (MP) [34] and its orthogonal variation (OMP) [65], Orthonormal Least Squares (OLS) [25] [14], and even the recent Stage-wise OMP [39].

The basic idea is to solve a  $N \times m$  linear system  $BX = Y$  by iterative subtraction of a vector proportional to the most significant column of  $B$ . The cost of such algorithms is typically a  $O(kmN)$  with  $N$  is the vector length ( $Y \in \mathbb{R}^N$ ),  $m$  is the size of the approximation family and  $k$  the number of nonzero coefficients of  $X$  in this family, but it can be reduced if the column vectors  $\{B_j\}_{j=1..m}$  admit a fast scalar product (discrete Fourier basis, wavelet bases, local cosines ...). In a non-orthogonal setting, these methods are not guaranteed to find the exact coordinates of the solution, although they converge quickly to a solution with a few nonzero coefficients in the family  $\{B_j\}_{j=1..m}$ . OMP and OLS have a larger numerical cost per iteration but converge faster to the solution.

Another set of methods called Basis Pursuit (BP) [26] or Frame Pursuit (FP) appeared which consist in the resolution a  $\ell^1$  constrained problem :

$$\min_u \|u\|_1 \text{ s.t. } \Phi \Gamma u = b \quad (\text{BP1})$$

$$\min_u \|u\|_1 \text{ s.t. } \|\Phi \Gamma u - b\|_2^2 \leq \epsilon \quad (\text{BP2})$$

aimed at recovering signals with a few nonzero coefficients  $u$  in the family of vectors  $(\Gamma_k)$ . The matrix  $\Phi$  is the measurement matrix and the number of lines of  $\Phi$  can be much smaller than the number of columns (the signal  $\Gamma u$  is undersampled in the basis formed by the row vectors  $(\Phi)_{(k,\cdot)}$  of  $\Phi$ ).

For the last ten years this problem has gathered people from the Signal Processing, Optimization and Statistics community and is known as *Compressive Sensing* (CS). The first algorithms to solve (BP1) were based on Linear Programming (LP) such as Interior Point methods, but researchers quickly noticed that this kind of problems could be solved by fast iterative shrinkage algorithms [33] [43] [44]. Works of Donoho et al. [38], Candes et al. [18] [19] and many others have contributed to a better understanding of the circumstances when the resolution of the Basis Pursuit problem (BP2) produces the right sparse solution and is stable to the presence of noise in the data.

In the following study of the microvibrations problem we use the ability of the  $\ell^1$ -norm to produce sparse vectors as observed in the literature, because the microvibrations of the push-broom devices can be well approximated by a sum of few pure sines. The way we conceived the following modelizations is closer to the original Basis Pursuit than Compressive Sensing. This latest concerns more specific problems for which a condition called *Restricted Isometry Property* (RIP) is verified, this is not necessarily the case in our study. Furthermore the disparities suffer from non-stochastic perturbations (elevation of buildings and ground) with large amplitudes (in the provided PLEIADES images, we have typically  $d(k) \in [-5, 5]$  and  $V_i(k) \in [-0.5, 0.5]$ ) and the sampling set depends on the perturbations, which is quite unusual.

### 3.2.2 Approximation with trigonometric polynomials

Because of the irregularity in the sampling grid of  $V_2$  (3.6), we have to consider continuous functions on  $\Omega = [0, m] \times [0, n]$ . The shape of the microvibrations (as described in [59]) is very close to a sum

of sine functions with specific frequencies, this is why we approximate them with the classical 1D trigonometric polynomials.

$$V_i(x) = \sum_{l=1-T}^{T-1} A_i(l) e^{il\frac{\pi}{T}x} \quad , \quad i \in \{1, 2\} \quad (3.7)$$

where the  $A_i$  are the Fourier coefficients and  $T$  is an integer parameter such that  $V_1$  and  $V_2$  are  $2T$ -periodic. The signals  $V_1$  and  $V_2$  are real so we impose Hermitian symmetry of the coefficients  $A_1(l)$  and  $A_2(l)$ , and since we use FFT on  $2T$  periodic signals we also impose one null coefficient at index  $T$  [2T].

$$\begin{cases} A_i(-l) &= \overline{A_i(l)} & 0 \leq l \leq T-1 \\ A_i(T) &= 0 \end{cases} \quad , \quad i \in \{1, 2\} \quad (3.8)$$

This approximation provides a discretization of the frequency interval  $[-\frac{1}{2}, \frac{1}{2}]$  at points  $\{\frac{l}{2T}\}_{1-T \leq l \leq T-1}$ . Indeed, by Landau's theorem [58] this is the largest bandlimited space with frequency interval centered on the origin we can retrieve with a sampling density  $dens(\Lambda) = 1$  (perturbed sampling may allow larger frequency intervals, but here the interval  $[-\frac{1}{2}, \frac{1}{2}]$  is enough). The frequency accuracy of this approximation is  $\frac{1}{2T}$  but a pure sine function with frequency  $\omega \notin \{\frac{l}{2T}\}_{0 \leq l \leq T-1}$  will generally be approximated by a sum of two or more (but few) sine functions with frequency belonging to the approximation grid.

$V_1$  is sampled on  $\Lambda$  with integer coordinates and  $V_2$  has sampling grid  $\Lambda'$  with non-integer coordinates. The first vector  $V_1$  can therefore be computed very quickly from  $A_1$  by discrete FFT.

The second microvibration  $V_2$  does not benefit from this fast implementation. We introduced in chapter 2 the sampling operator of a trigonometric polynomial and detailed the properties of such operators. The sampling matrix  $L_2$  is a complex Vandermonde matrix with general term

$$(L_2)_{k,l} = e^{i\frac{\pi}{T}l x'_k} \quad \text{for all } 1-T \leq l \leq T-1 \quad 0 \leq k \leq |\Lambda|$$

where  $x'_k = x_k - (d_k)_1$ .

The operators  $L_1^* L_1$  and  $L_2^* L_2$  are intensively used in our algorithms, it is thus crucial to implement the matrix-vector multiplication with these operators in an efficient way. The lines of  $L_1$  are simply repeated lines of the classical discrete Fourier transform matrix (with possible gaps) :

$$(L_1 A_1)_j = (\mathcal{F}(A_1))_{x_j [2T]}$$

and the product  $L_1^* L_1 A_1$  can be computed by means of simple algebra operations and discrete FFT. This operation thus has a cost of approximately 2 FFTs on  $2T$ -length signals.

$$(L_1^* L_1) A_1 \asymp O(2T \log 2T)$$

Computation of the product  $L_2^* L_2 A_2$  necessarily uses the Toeplitz property of  $L_2^* L_2$  because of the irregularity of the grid  $\Lambda_2$ . We remind that  $P_2 = L_2^* L_2$  is a Toeplitz matrix verifying

$$(P_2)_{k,l} = \sum_{j=1}^{N_\Lambda} e^{i\frac{\pi}{T}(k-l)x_j}$$

and the computation of matrix-vector product benefits from the FFT :

$$\begin{aligned} (P_2 A_2)_k &= (p_2 * \tilde{A}_2) [k] \\ &= (\mathcal{F}^{-1}(\mathcal{F}(p_2) \times \mathcal{F}(\tilde{A}_2))) [k] \quad 1-T \leq k \leq T-1 \end{aligned}$$

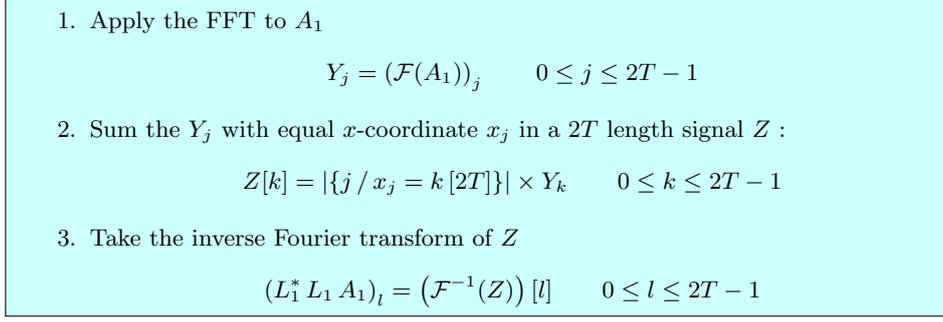


Figure 3.4: Fast computation of the matrix-vector product  $L_1^* L_1 A_1$  where  $L_1$  is the sampling matrix of a trigonometric polynomial with equal period and degree on a grid  $\Lambda \subset \mathbb{N}$ .

where  $p_2$  and  $\tilde{A}_2$  are the  $4T$ -length signals defined by :

$$p_2[l] = \begin{cases} (P_2)_{j,k} & \text{s.t. } k - j = l [4T] \\ 0 & \text{else} \end{cases} \quad \text{if } l \neq 2T[4T]$$

$$\tilde{A}_2[l] = \begin{cases} A_2(l) & \text{for } 1 - T \leq l \leq T - 1 \\ 0 & \text{else} \end{cases}$$

This requires the precomputation of  $\hat{p}_2$  (accounts for a non-negligible part of the computation time of algorithms presented in this section) and the numerical cost is finally :

$$\begin{aligned} L_2 A_2 &\asymp O(T |\Lambda|) \\ \mathcal{F}(p_2) &\asymp O(2T |\Lambda|) + O(4T \log 4T) \quad (\text{initialization}) \\ P_2 A_2 &\asymp O(4T \log 4T) \quad (\text{for each multiplication}) \end{aligned} \tag{3.9}$$

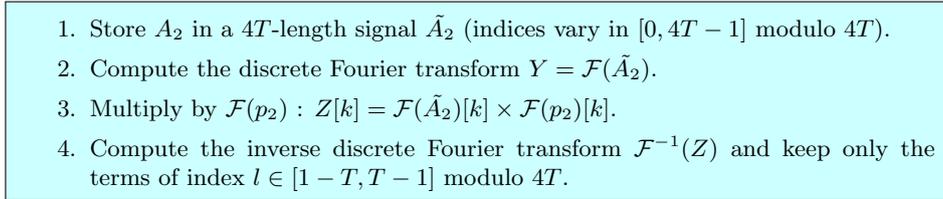


Figure 3.5: Fast computation of the matrix-vector product  $L_2^* L_2 A_2$  where  $L_2$  is the sampling matrix of a trigonometric polynomial.

The main drawback of trigonometric polynomials is that  $L_2$  has no fast implementation because of the irregularity of  $\Lambda'$ . The resolution of the different problems we address in the next section was influenced by this numerical constraint and we tried, whenever it was possible, to use a fast implementation of previous matrices. A numerically satisfying solution to decrease the computation cost of matrix-vector products  $L_2 A_2$  would be the use of the NFFT (*Non-uniform Fast Fourier Transform*) of Keiner et al. [56] with the drawback of small numerical errors (but it is possible to make this approximation error arbitrarily small by refining the approximation grid of the NFFT). This would decrease the computation cost to a  $O(T \log T) + O(|\Lambda|)$  (the term in  $T \log T$  is larger

than the usual FFT) and should be very helpful for large data sets. In the case of detection and estimation algorithms below, the computation times for  $600 \times 600$  images was reasonable (about one minute). In section 3.4 we propose algorithms slower algorithms (about hours) but the main part of this computation time is not due to the  $L_2$

### 3.2.3 Variational formulations of the problem

#### 3.2.3.1 General formulation

In our problem the available data is made of a disparity map and a mask which determines the sampling grid  $\Lambda$ . The distribution of available points can be very chaotic as the MARC2 software eliminates several kind of disadvantages from the patch comparison method (see figure 3.6).

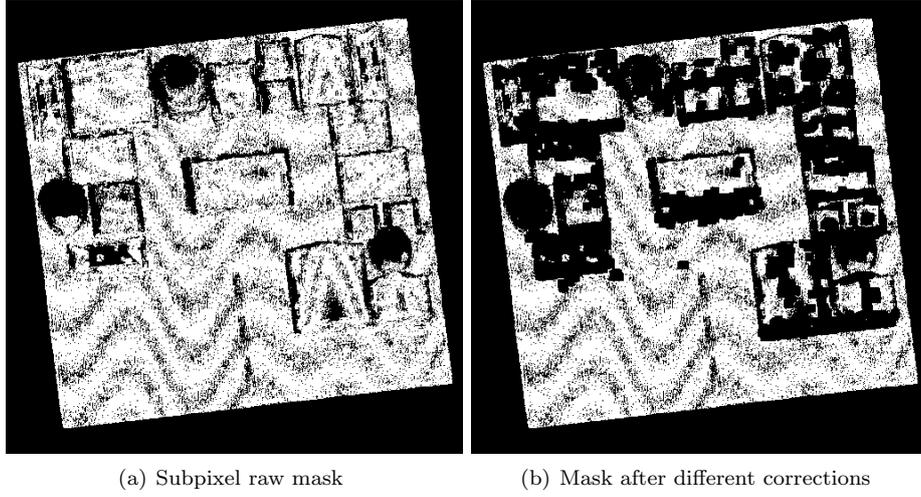


Figure 3.6: The different output masks in MARC2 : correction of adhesion and false alarms can produce very complicated maps. (some roofs are not detected because of the chosen disparity range  $[-5, 5]$ ).

Tools like selective FFT filtering or other filtering methods use a full density map and introduce thresholds, here we propose to determine the signals  $V_1$  and  $V_2$  from lacunary density maps and a single parameter : the number of frequencies with nonzero amplitudes (or the  $l^0$  value of the Fourier coefficients). In the spirit of Basis Pursuit, we consider the following minimization problem ( $\mathcal{P}_1$ ) and its Lagrangian formulation.

$$\begin{aligned}
 (A_1, A_2) &= \underset{(A, B)}{\operatorname{argmin}} E(A, B) \text{ s.t. } \|A\|_1 \leq C_1 \text{ and } \|B\|_1 \leq C_2 \\
 \Rightarrow \exists \lambda_1, \lambda_2 > 0 \text{ s.t. } (A_1, A_2) &= \underset{(A, B)}{\operatorname{argmin}} E(A, B) + \lambda_1 \|A\|_1 + \lambda_2 \|B\|_1
 \end{aligned}
 \tag{\mathcal{P}_1}$$

where  $E(A, B) = \frac{1}{2} \|D + L_1 A - L_2 A\|_2^2 = \frac{1}{2} \left\| D - L \begin{pmatrix} A \\ B \end{pmatrix} \right\|_2^2$  and  $L = (-L_1, L_2)$ .

### 3.2.3.2 Stationary case

As we considered stationary and non-stationary signals, we can add a stationarity constraint to  $(\mathcal{P}_1)$  or impose

$$V_2(t) = V_1(t + \tau) \quad \Leftrightarrow \quad A_2 = S_\tau A_1$$

$$\text{with } (S_\tau)_{k,l} = \delta_{k,l} e^{i\frac{2\pi}{T}l\tau}$$

This leads to the two following problems

$$\boxed{(A_1, A_2) = \underset{(A,B)}{\operatorname{argmin}} E(A, B) + \lambda_1 \|A\|_1 + \lambda_2 \|B\|_1 + \frac{\lambda_3}{2} \|S_\tau A_1 - A_2\|_2^2} \quad (\mathcal{P}_2)$$

$$\boxed{A_1 = \underset{A}{\operatorname{argmin}} E(A, S_\tau A) + (\lambda_1 + \lambda_2) \|A\|_1} \quad (\mathcal{P}_3)$$

In  $(\mathcal{P}_2)$  we added a quadratic constraint on the Fourier coefficients of  $V_1$  and  $V_2$  and in  $(\mathcal{P}_3)$  we impose the exact time shift between the two microvibrations. This formulation preserves the  $l^1 + (l^2)^2$  structure in  $(\mathcal{P}_1)$  and could be implemented with very few changes. Problem  $(\mathcal{P}_3)$  is simpler since the number of parameters is halved compared to  $(\mathcal{P}_1)$  and  $(\mathcal{P}_2)$ .

## 3.2.4 Resolution of problem $(\mathcal{P}_1)$ and its variations

### 3.2.4.1 Minimization with the FB/ISTA algorithm

Let  $E_1(A, B) = E(A, B) + \lambda_1 \|A\|_1 + \lambda_2 \|B\|_1$  be the minimized functional in problem  $(\mathcal{P}_1)$ . This function is a the sum of two convex continuous functions with  $E(A, B)$  a differentiable function with affine gradient.

$$\begin{aligned} \nabla E(A, B) &= L^* \left( L \begin{pmatrix} A \\ B \end{pmatrix} - D \right) \\ &= \begin{pmatrix} L_1^* L_1 & -L_1^* L_2 \\ -L_2^* L_1 & L_2^* L_2 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} - \begin{pmatrix} -L_1^* D \\ L_2^* D \end{pmatrix} \end{aligned} \quad (3.10)$$

so we can solve problem  $(\mathcal{P}_1)$  by the semi-implicit ISTA algorithm 3.7 :

0. Initialize  $n = 0$ ,  $A^0$ ,  $B^0$  and  $l > 0$  the Lipschitz constant of  $L^* L$  (3.10) and  $\rho < \frac{2}{l}$ . Compute  $L_1^* D$  and  $L_2^* D$ .
  1. **Repeat**
    - (a)  $\begin{pmatrix} A \\ B \end{pmatrix}^{n+1/2} = \begin{pmatrix} A \\ B \end{pmatrix}^n - \rho \nabla E(A^n, B^n)$
    - (b)  $(A, B)^{n+1} = \operatorname{prox}_{\rho\lambda_1 \|A\|_1 + \rho\lambda_2 \|B\|_1}(A^{n+1/2}, B^{n+1/2})$   
 $= \operatorname{soft}_{(\rho\lambda_1, \rho\lambda_2)}(A^{n+1/2}, B^{n+1/2})$
    - (c)  $n \leftarrow n + 1$

**while**  $|E_1(A^n, B^n) - E_1(A^{n-1}, B^{n-1})| > \epsilon E_1(A^{n-1}, B^{n-1})$  and  $n < n_{max}$ .
  2.  $V_1 = L_1 A^n$ ,  $V_2 = L_2 B^n$  and  $D^* = D + V_1 - V_2$

Figure 3.7: Forward-Backward algorithm applied to problem  $(\mathcal{P}_1)$ . The corrected disparity vector here is denoted by  $D^*$ .

We also proposed an alternate minimization scheme in order to compare the gain of fast implementations of Toeplitz matrices above and full-matrices multiplications. Indeed, in equation (3.10) we note the presence of two non-Toeplitz matrices  $L_1^* L_2$  and  $L_2^* L_1$  which penalizes the computation of matrix-vector products. Without explicit computation of the matrix  $L_1^* L_2$  we have

$$L_1^*(L_2 B) \asymp L_2^*(L_1 A) \asymp O(2T(\log 2T + |\Lambda|))$$

The cost of each matrix-vector multiplication can be reduced to  $O(T^2)$  if this matrix is explicitly computed and stored :

$$\begin{aligned} (L_1^* L_2) &\asymp O(2T(\log 2T + |\Lambda|)) && \text{(initialization)} \\ \Rightarrow (L_1^* L_2) A_2 &\asymp (L_2^* L_1 A_1) \asymp O(T^2) && \text{(for each multiplication)} \end{aligned}$$

whereas the minimization in  $A$  or  $B$  uses only Toeplitz matrices but requires an update of the constant part of the gradient in (3.10) at each iteration.

0. Set  $n = 0$ , initialize  $A^0$ ,  $B^0$ . Let  $\rho < \frac{2}{7}$  with  $l > 0$  the maximum Lipschitz constant among  $L_1^* L_1$  and  $L_2^* L_2$ . Compute  $L_1^* D$  and  $L_2^* D$ .

1. **Repeat**

(a) Set  $k = 0$ ,  $\tilde{A}^0 = A^n$ . Compute  $L_1^* L_2 B^n$ .

(b) **Repeat**

- $\tilde{A}^{k+1/2} = \tilde{A}^k - \rho (L_1^* L_1 \tilde{A}^k - L_1^* L_2 B^n - L_1^* D)$
- $\tilde{A}^{k+1} = \text{soft}_{\rho \lambda_1}(\tilde{A}^{k+1/2})$
- $k \leftarrow k + 1$

**while**  $|E_1(\tilde{A}^k, B^n) - E_1(\tilde{A}^{k-1}, B^n)| > \nu E_1(\tilde{A}^{k-1}, B^n)$  and  $k < k_{max}$ .

(c)  $A^{n+1} = \tilde{A}^k$

(d) Set  $k = 0$ ,  $\tilde{B}^0 = B^n$ . Compute  $L_2^* L_1 A^{n+1}$ .

(e) **Repeat**

- $\tilde{B}^{k+1/2} = \tilde{B}^k - \rho (L_2^* L_2 \tilde{B}^k - L_2^* L_1 A^{n+1} + L_2^* D)$
- $\tilde{B}^{k+1} = \text{soft}_{\rho \lambda_2}(\tilde{B}^{k+1/2})$
- $k \leftarrow k + 1$

**while**  $|E_1(A^{n+1}, \tilde{B}^k) - E_1(A^{n+1}, \tilde{B}^{k-1})| > \nu E_1(A^{n+1}, \tilde{B}^{k-1})$  and  $k < k_{max}$ .

(f)  $B^{n+1} = \tilde{B}^k$

(g)  $n \leftarrow n + 1$

**while**  $|E_1(A^n, B^n) - E_1(A^{n-1}, B^{n-1})| > \epsilon E_1(A^{n-1}, B^{n-1})$  and  $n < n_{max}$ .

2.  $V_1 = L_1 A^n$ ,  $V_2 = L_2 B^n$  and  $D^* = D + V_1 - V_2$

Figure 3.8: Alternate minimization of problem  $(\mathcal{P}_1)$ . Each sub-problem involves only Toeplitz matrices except for the initialization and has very fast convergence.

$$(L_1^* L_1) A \asymp O(2T \log 2T)$$

$$(L_2^* L_2) B \asymp O(4T \log 4T)$$

Alternate minimization 3.8 consists here in inner loops with fast convergence due to the Toeplitz matrices but the outer loop includes the same kind of computations as one iteration in 3.7. The alternate minimization provided early results which encouraged us to implement the full Forward-Backward algorithm 3.7 with non-Toeplitz operators. Problem  $(\mathcal{P}_2)$  is easy to implement from the previous algorithm as the additional quadratic term  $\lambda_3 \|S_\tau A_1 - A_2\|_2^2$  in energy  $E(A, B)$  becomes a linear term in  $\nabla E(A, B)$  (3.11).

$$\nabla \left( \frac{\lambda_3}{2} \|S_\tau A_1 - A_2\|_2^2 \right) = \lambda_3 \begin{pmatrix} A_1 - S_{-\tau} A_2 \\ A_2 - S_\tau A_1 \end{pmatrix} \quad (3.11)$$

since  $S_\tau^{-1} = S_{-\tau}$ . In  $(\mathcal{P}_2)$  the Lipschitz constant  $l$  of  $L$  is replaced with the Lipschitz constant of  $L + \lambda_3 P$  (see algorithm 3.17 in section 3.2.6), and should be chosen approximately equal to  $\max(l, \lambda_3)$ . This is clearly a drawback because the convergence speed of the Forward-Backward algorithm depends on this Lipschitz constant and its estimate.

### 3.2.4.2 Detection results

The first step in our study consists in the detection of the right frequencies in the disparity map. Tables 3.0(a) and 3.0(b) contain the different parameters of the microvibrations in the test images we used. The two main frequencies are  $0.0048672 \text{ pix}^{-1}$  and  $0.008112 \text{ pix}^{-1}$  in both cases, but in the Village images we reduced the number of nonzero frequencies to the number of principal modes of PLEIADES. Indeed, our experiments on the PLEIADES couples showed that the other modes are never detected by our method. The shift parameters of each frequency in table 3.0(b) correspond to the time/space values  $t_{j,l}$  in the following expression :

$$\begin{aligned} V_j(\cdot) &= \sum_{l=0}^{T-1} |A_j(l)| e^{i \frac{\pi}{T} l (\cdot - t_{j,l})} \\ \Leftrightarrow \arg(A_j(l)) &= -\frac{\pi}{T} l t_{j,l} \quad [2\pi] \end{aligned} \quad (3.12)$$

(a) PLEIADES images (unknown phase ratios in the microvibrations)

Frequencies	Amplitudes
0,0048672	0,25
0,008112	0,25
0,0356928	0,075
0,0397488	0,082
0,04056	0,082
0,056784	0,041

(b) Village images (imposed time ratio  $\tau = 75.91$ )

Frequencies	Amplitudes	Shift ( $V_1$ )	Shift ( $V_2$ )
0,0048672	0,25	0	75,91
0,008112	0,25	34,12	110,03

Table 3.1: Parameters of the microvibrations in the different image couples of this test. Microvibrations in PLEIADES images have six nonzero frequencies with unknown phases (or shift) and the Village couple microvibrations contain only two nonzero frequencies with known phases (the time ratio between  $V_1$  and  $V_2$  is known and here equal to 75.91)

The simulated PLEIADES images provided by the CNES<sup>3</sup> are available in two versions : i) a

<sup>3</sup>We thank the CNES and specially G. Blanchet for providing these images.

first version in which one frequency is in phase opposition in the two microvibrations ii) a second version with random phases microvibrations. In both cases the phases or shift of each frequency are unknown, consequently the only valid model for these images is problem ( $\mathcal{P}_1$ ). The second version (random phases) should be harder to restore because a mixing between the two microvibrations has more chances to occur. Each couple of coefficients ( $A_1(l)$ ,  $A_2(l)$ ) will add a sine looking perturbation to the disparity  $D$  as shown in figure 3.3 and in regions of constant disparity  $H$  this perturbation will be close to :

$$2|A(l)| \sin\left(\frac{\phi_2(l) - \phi_1(l)}{2}\right) \sin\left(\omega(l)x + \frac{\phi_1(l) + \phi_2(l)}{2}\right)$$

where  $\omega(l)$  is the corresponding frequency and  $\phi_1(l)$ ,  $\phi_2(l)$  the arguments of  $A_1(l)$  and  $A_2(l)$ . Phase opposition for a given frequency means that  $\phi_1(l) - \phi_2(l) = \pi [2\pi]$  and the associated amplitude is thus twice larger than  $|A_1(l)|$ . Stationary microvibrations such that  $V_2(t) = V_1(t + \tau)$  cannot in general have all frequencies in phase opposition since the amplitude associated to frequency  $\omega(l)$  in the difference ( $V_1 - V_2$ ) is

$$2|A_1(l)| \sin\left(\frac{\omega(l)\tau}{2}\right).$$

The relation  $\omega(l)\tau \in \pi [2\pi]$  between the oscillation modes of the satellite (the  $\omega(l)$ s) and the shift  $\tau$  is certainly never verified for every frequency mode in table 3.0(a).

In all the cases we considered<sup>4</sup>, the amplitudes of each frequency in both microvibrations are equal so the Lagrange parameter  $\lambda_1$  and  $\lambda_2$  should be very close. In figures 3.10 and 3.11 we display the evolution of the  $\ell^0$  and  $\ell^1$  values of the solution as a function of the Lagrange parameters (with  $\lambda_1 = \lambda_2$ ,  $\lambda_2 \in [2000, 15000]$ ). The moduli of the coefficients  $A_1(l)$  and  $A_2(l)$  in figures 3.12 and 3.13 show that in general the opposition case is indeed better detected, as the two principal modes  $\approx 0.0048 \text{ pix}^{-1}$  and  $\approx 0.0081 \text{ pix}^{-1}$  have similar amplitudes. These amplitudes can be very different in the random case (this happens in all cases of figure 3.13) which means that the target frequency has the same phase in both microvibrations and thus a lower amplitude in the disparity map. In the general case, and up to the convergence precision, we observe that the number of nonzero coefficients ( $\ell^0$  value) in each vector  $A_1$  and  $A_2$  is a piecewise constant, decreasing function when the Lagrange parameters  $\lambda_1$  and  $\lambda_2$  increase. We verify experimentally that the sum of the  $\ell^1$  norms is a strictly decreasing function, but this is not the case for each vector separately (figure 3.11 (top)).

---

<sup>4</sup>The study of completely independent microvibrations is possible with model ( $\mathcal{P}_1$ ), but in practise this case is not verified.

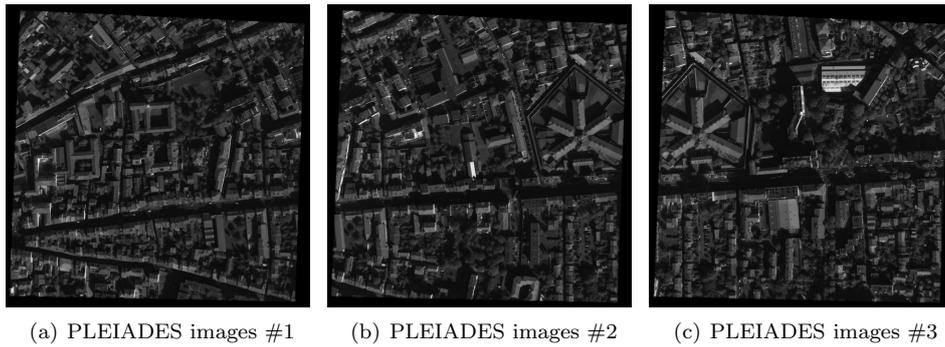


Figure 3.9: Test images (CNES copyrighted simulated PLEIADES images of Toulouse with resolution 0.7m). The couples have been rectified to epipolar geometry before the addition of horizontal microvibrations.

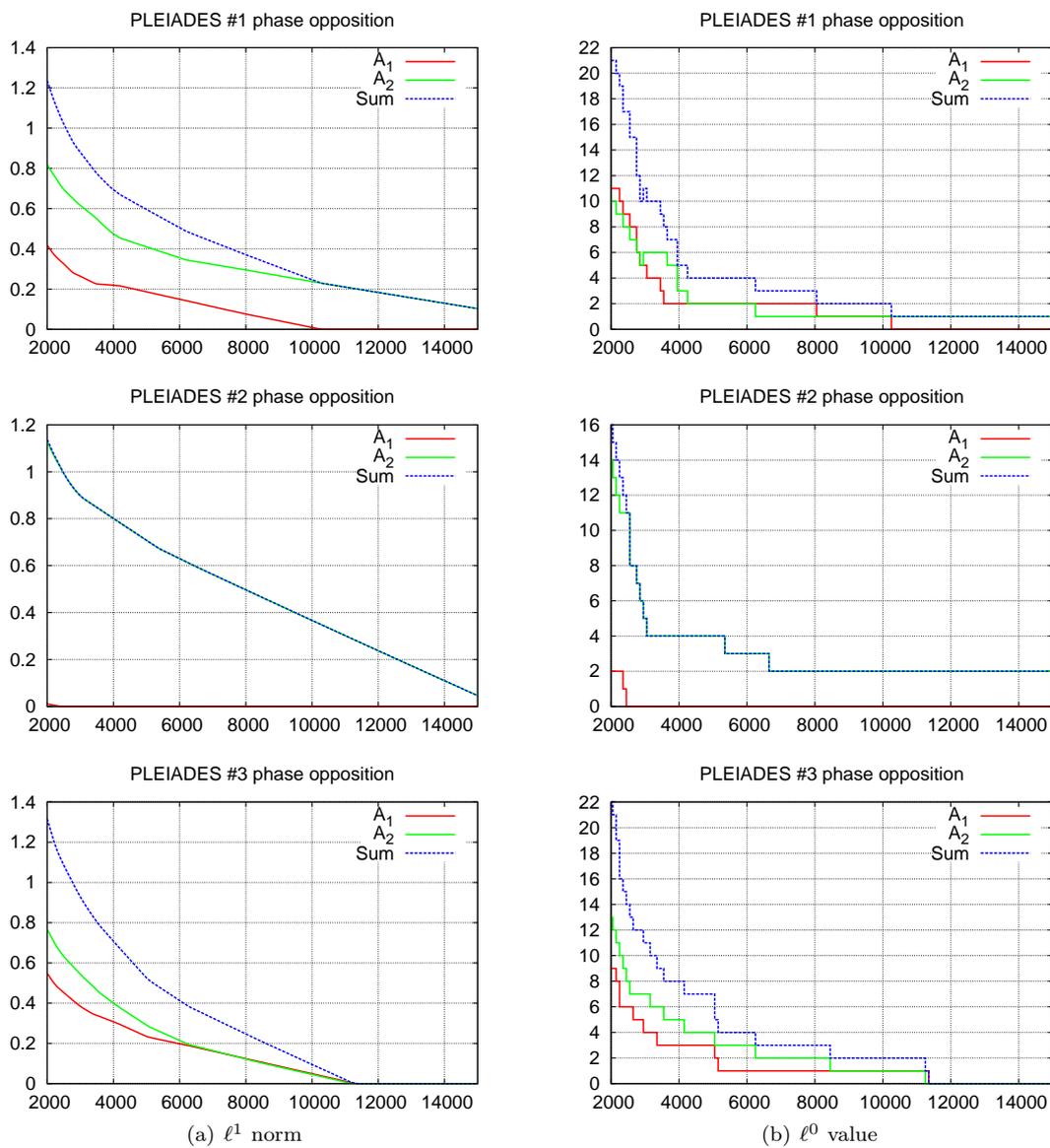


Figure 3.10:  $\ell^1$  norms and  $\ell^0$  values of both  $A_1$ ,  $A_2$  and sum as a function of the Lagrange parameters  $\lambda_1$  and  $\lambda_2$  with  $\lambda_1 = \lambda_2$ : phase opposition (one of the frequencies is in phase opposition). The left images are displayed in figure 3.9.

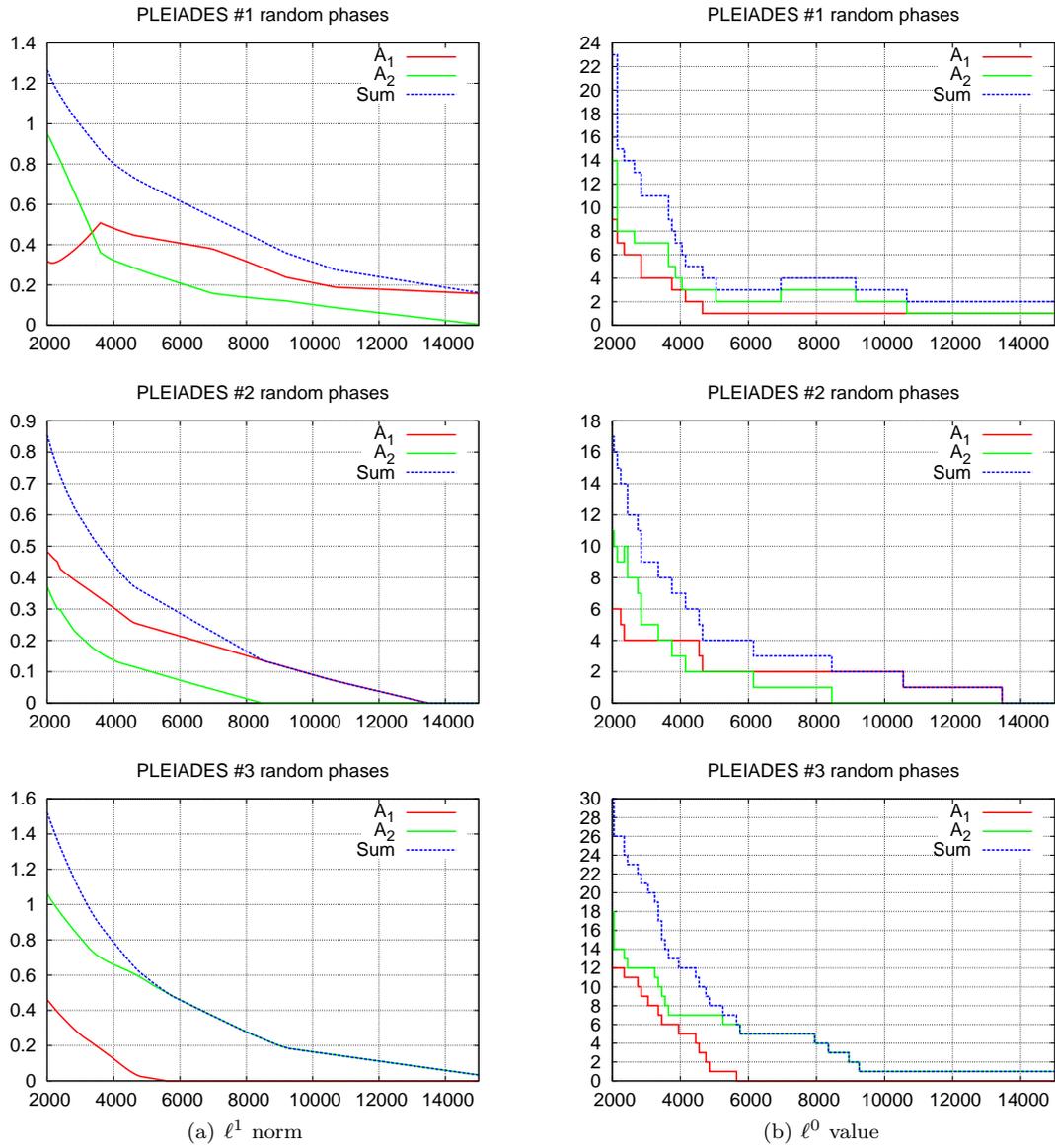


Figure 3.11:  $\ell^1$  norms and  $\ell^0$  values of both  $A_1$ ,  $A_2$  and sum as a function of the Lagrange parameters  $\lambda_1$  and  $\lambda_2$  with  $\lambda_1 = \lambda_2$  : random phases.

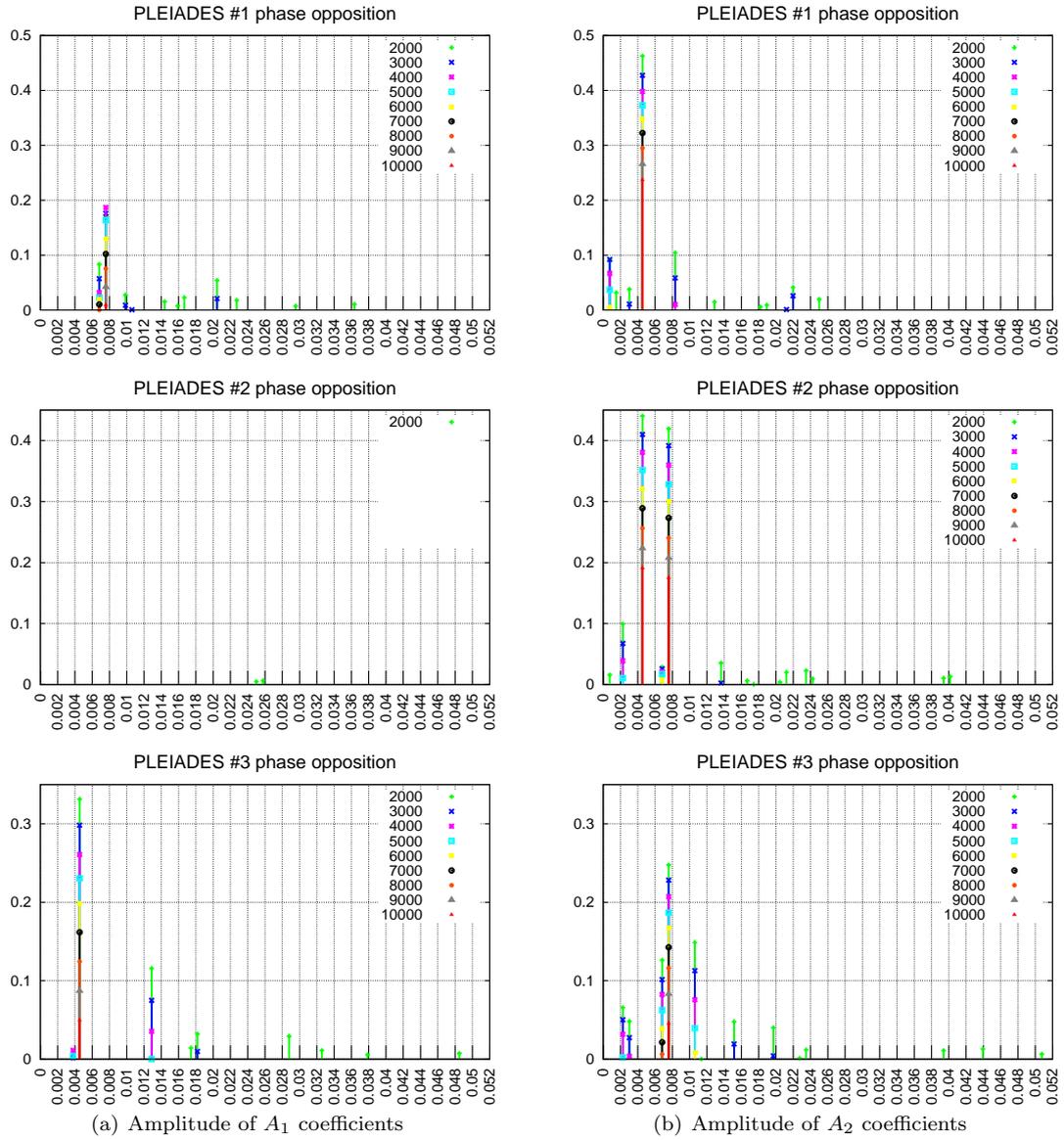


Figure 3.12: Amplitudes of coefficients in the solutions  $A_1$  and  $A_2$  : phase opposition (Frequencies in  $\text{pix}^{-1}$  on horizontal axis) . Our experiments tend to prove that the frequency in phase opposition is 0.0048672 although its amplitude in the third row is lower than the expected 0.5

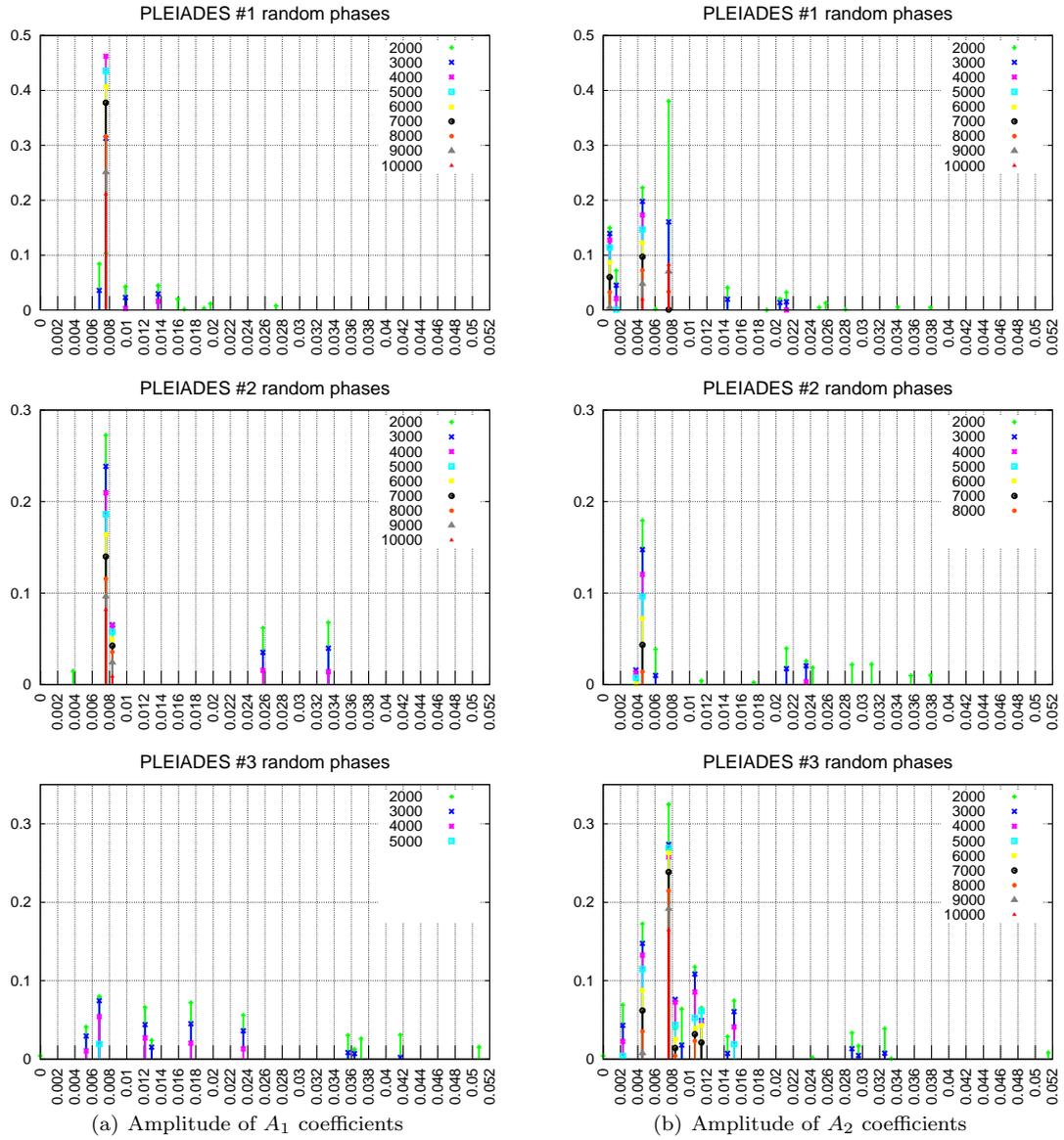


Figure 3.13: Amplitudes of coefficients in the solutions  $A_1$  and  $A_2$  : random phases. This clearly has an influence in the third case where the inlier 0.0045 is hardly detected compared to figure 3.12.

Instinctively (and for practical reasons) we chose to explore the Lagrange domain with  $\lambda_1 = \lambda_2$ , but we observe that in general it produces vectors  $A_1$  and  $A_2$  with different  $\ell^0$  values. Tuning these parameters in such a way that  $A_1$  and  $A_2$  have the same number of nonzero coefficients is still possible with an outer Uzawa loop, but the computation time can become arbitrarily large in this case.

Indeed, the update of one parameter can decrease the  $\ell^0$  value of one vector and increase the other. This phenomenon is not a critical point in the detection step since we do not take care about the associated amplitudes, but it is probable that the amplitudes estimation will suffer from this. We observe on the one hand that irregularity in the sampling of  $V_2$  produces different vectors  $A_1$  and  $A_2$ , and on the other hand the separation of these two microvibrations will be a difficult task because of the mixing between these two vectors. This is illustrated in figure 3.14 where we chose Lagrange parameters  $\lambda_1$  and  $\lambda_2$  such that  $\lambda_1 + \lambda_2 = C^{te}$ . We observe the approximative conservation of the quantity  $|A_1(l)| + |A_2(l)|$  for the principal frequencies, which means that a portion of the disparity can be explained with either  $V_1$  or  $V_2$ .

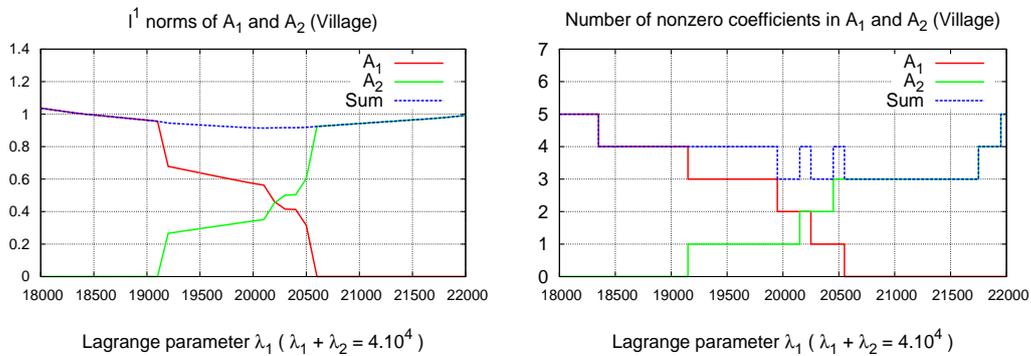


Figure 3.14: Mixing between  $V_1$  and  $V_2$  : increasing the penalty on one vector transfers the amplitudes to the other. This graph also shows how to choose parameters  $\lambda_1$  and  $\lambda_2$  such that  $\|A_1\|_1 = \|A_2\|_1$  (here 20200 and 19800). In this case the  $\ell^0$  values of these vectors are the same ( $\ell^0(A_1) = \ell^0(A_2) = 2$ ).

These first tests on PLEIADES simulated images seem to detect correctly the principal modes in the microvibrations but it is not possible here to check the accuracy of the correction since we do not know the ground truth. In the next experiment we applied this detection algorithm to the image Village from L. Moisan's simulator (figure 3.3(c) p. 38).

### 3.2.5 Partial correction

These results can be used to remove a part of the microvibrations in the disparity map despite the difficulty in the tuning of Lagrange parameters. The partially corrected map can be viewed in figure 3.16. The data Village does not satisfy the decorrelation hypothesis between the true disparity signal  $H(k + V_1(k))$  and the microvibrations despite the presence of a large plane area. The spacing between buildings and their resolution are certainly at the origin of this mixing and illustrates the case of correlated microvibrations and disparity.

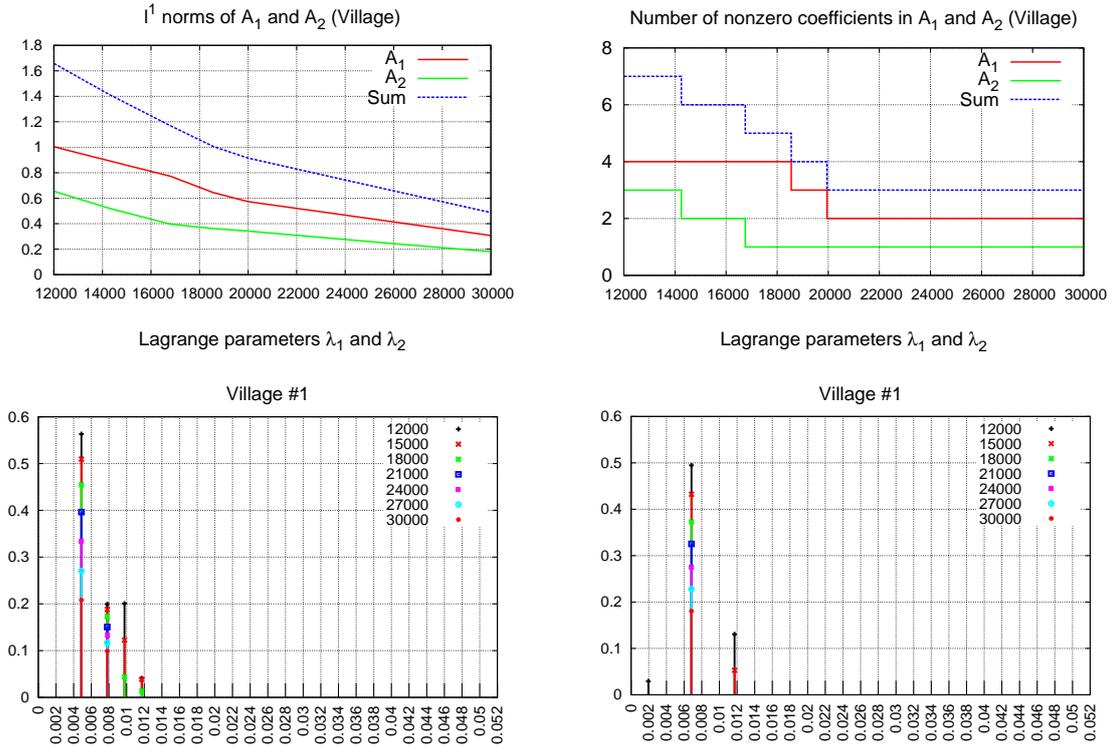


Figure 3.15: Detection algorithm applied to the simulated Village images. The frequency 0.008 is indeed detected but the associated amplitude decreases faster than the 0.0068 outlier, partial correction is not accurate at all (3.16).

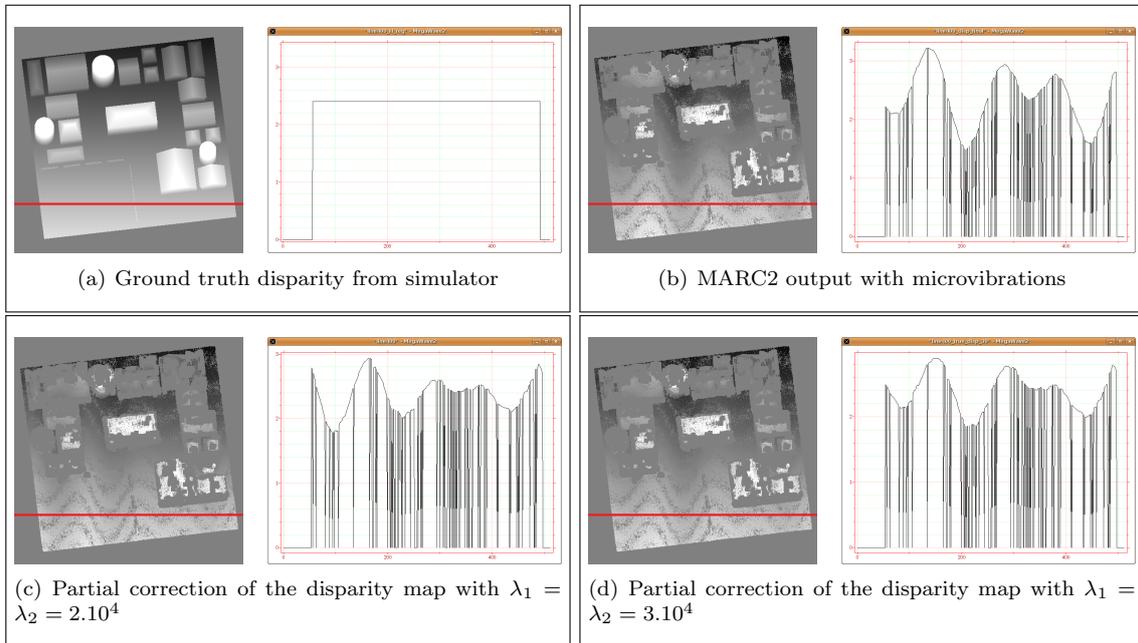


Figure 3.16: Partial correction of the disparity map after the detection step (here with large parameters  $\lambda_1 = \lambda_2 = 3.10^4$ ). Despite the large plane areas in the ground truth the detection contains an outlier and does not correctly estimate the perturbation.

The next section is a resolution of problem  $(\mathcal{P}_3)$  applied to the Village images. We imposed a known time shift between the two microvibrations of this couple and propose to reduce the number of unknowns to check whether this mixing phenomenon persists or not with additional knowledge on the oscillations.

### 3.2.6 Stationary microvibrations

Our solving of the stationary case was achieved first by a quadratic penalization on the difference  $S_\tau A_1 - A_2$  where  $S_\tau$  is a diagonal matrix with general term  $(S_\tau)_{l,l} = e^{i\frac{\pi}{T}l\tau}$ . It corresponds to the diagonalization of the shift operator acting on the set of trigonometric polynomials :

$$\begin{aligned} \sum_{l=1-T}^{T-1} A_1(l) e^{i\frac{\pi}{T}l(+\tau)} &= \sum_{l=1-T}^{T-1} (A_1(l) e^{i\frac{\pi}{T}l\tau}) e^{i\frac{\pi}{T}l} \\ &= \sum_{l=1-T}^{T-1} (S_\tau A_1)_l e^{i\frac{\pi}{T}l}. \end{aligned}$$

This problem suffers from the following disadvantage : if one wants to impose exact time shift between  $V_1$  and  $V_2$ , the associated Lagrange  $\lambda_3$  in  $(\mathcal{P}_2)$  should be very large (a factor  $\approx 100$  between  $\lambda_1$  and  $\lambda_3$  was necessary in our case) and so would the Lipschitz constant of the associated linear operator. We mention in algorithm 3.17 the modifications to algorithm 3.7 corresponding to this implementation.

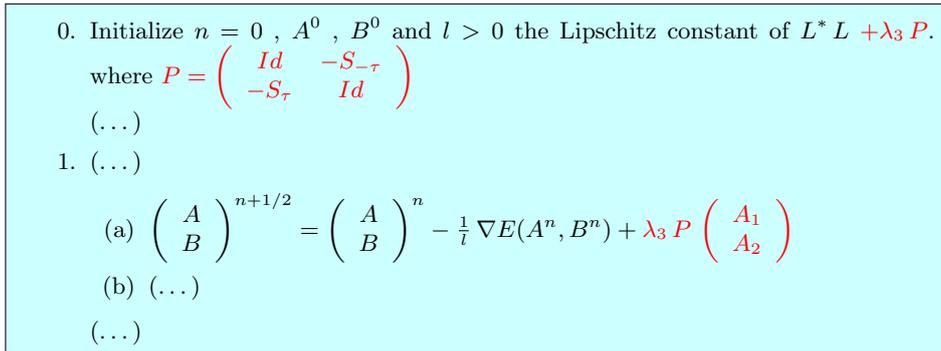


Figure 3.17: Modified Forward-Backward algorithm for  $(\mathcal{P}_2)$ .

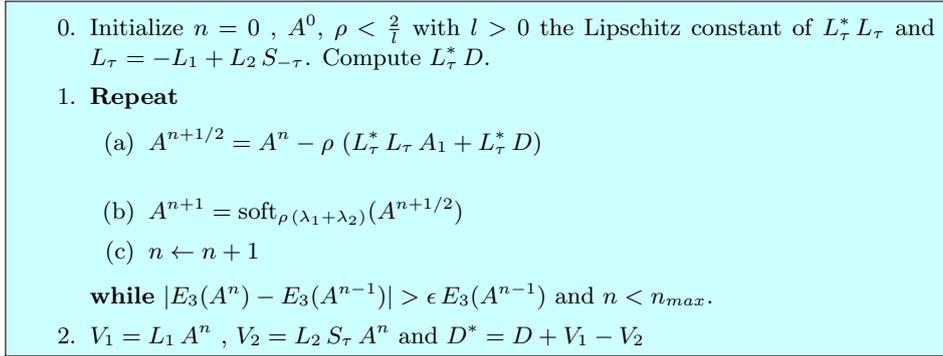
This formulation allows flexibility in the microvibrations, for example in the case when the stationarity hypothesis is not exactly verified, an  $l^2$  penalization between the two vectors  $A_1$  and  $A_2$  can allow slight differences in amplitudes and phases of  $S_\tau A_1$  and  $A_2$ . Concerning problem  $(\mathcal{P}_2)$ , we verified experimentally that for large parameters  $\lambda_3$  the solution of  $(\mathcal{P}_2)$  converges to the solution  $(\mathcal{P}_3)$  for the  $l^2$  norm (but the number of nonzero coefficients is generally larger with  $(\mathcal{P}_3)$ ).

We focus now on the simpler problem  $(\mathcal{P}_3)$ . The scheme of the algorithm is similar to 3.7 but gradient of the  $l^2$  part no more has Toeplitz structure :

$$\begin{aligned} \nabla \left( \frac{1}{2} \|(-L_1 + L_2 S_{-\tau}) \cdot -D\|^2 \right) (A_1) &= (L_1^* L_1 + (L_2 S_{-\tau})^* (L_2 S_{-\tau}) \\ &\quad - (L_2 S_{-\tau})^* L_1 - L_1^* L_2 S_{-\tau}) A_1 - (-L_1 + L_2 S_{-\tau})^* D \end{aligned}$$

and the algorithm should be adapted according to algorithm 3.18.

Experiments on the Village couple (figure 3.19) unfortunately shows the persistence of the outlier under stationarity hypothesis with algorithm 3.18. The resolution of the scene is thus

Figure 3.18: Forward-Backward algorithm for  $(\mathcal{P}_3)$ .

very important when using this blind separation approach and we observe that the stationarity hypothesis is not sufficient in this problem.

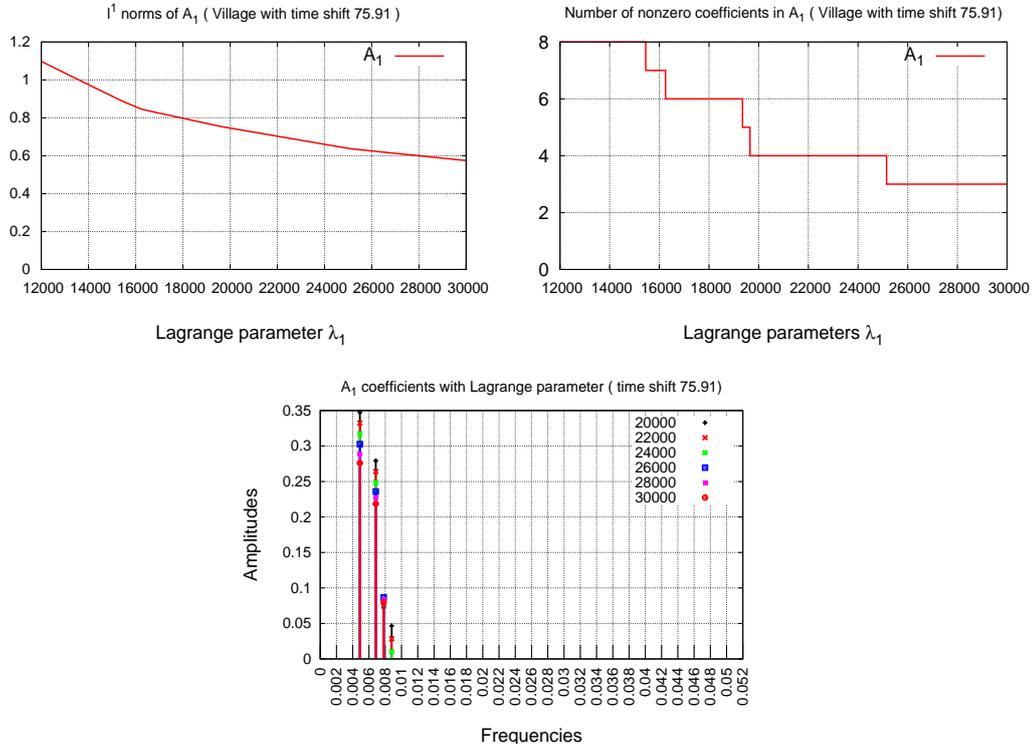


Figure 3.19: Detection of microvibrations with formulation  $(\mathcal{P}_3)$  and a time shift corresponding to the real time shift between  $V_1$  and  $V_2$ . The outlier  $0.0068 \text{ pix}^{-1}$  is still present and the correction is not improved compared to 3.16

### 3.3 Amplitude estimation

Results from the detection step show that it is necessary to over-regularize the problem in order to remove wrong detections (outliers). In this case correct detections suffer from the so-called *erosion* effect which occurs in every  $\ell^1$  constrained problem. This drawback of Basis Pursuit can be corrected by solving the least-squares problem with restriction to the detected frequencies provided that they do not contain outliers. We present in this section two approaches to perform this additional least-squares step.

#### 3.3.1 Iterated Hard-Thresholding Algorithm (IHTA)

Erosion is a well known phenomenon due to soft-thresholding and works by Bluemensath et al. [13] have shown that replacing this soft-thresholding by a hard-thresholding in the Forward-Backward algorithm gives a local minimizer of a  $l^0$  Lagrangian constrained function.

**Theorem 14 (Blumensath, Davies 2008)** *Let  $C_{L_0}(X) = \|\Phi X - Y\|^2 + \lambda l^0(X)$  with  $\lambda > 0$ ,  $\Phi$  a linear operator on  $\mathbb{R}^d$  such that  $Id - \Phi^t \Phi$  has a Lipschitz constant  $< 1$ . The sequence  $(X^k)_{k \geq 0}$  defined by  $X^0 \in \mathbb{R}^d$  and  $\forall k \geq 0$  :*

$$X^{k+1} = \text{hard}_{\sqrt{\lambda}}(X^k - \Phi^t(\Phi X^k - Y)) \quad (3.13)$$

*converges to a local minimizer of  $C_{L_0}$ .*

with

$$(\text{hard}_{\lambda}(X))_i = \begin{cases} 0 & \text{if } |X_i| \leq \lambda \\ X_i & \text{else} \end{cases} \quad (3.14)$$

By a rescaling argument we obtain that the hard-thresholding algorithm with threshold  $\rho\lambda$  and gradient descent step  $\rho$  as in (3.7) converges to a local minimizer of :

$$\frac{1}{2} \|\Phi X - Y\|^2 + \frac{\rho\lambda^2}{2} l^0(X) \quad (3.15)$$

The local minimizer is an optimal result because the  $\ell^0$  Lagrangian constrained problem in this case is a NP-hard problem (see [65] p. 612). It is thus necessary to initialize this algorithm close to the global minimizer to avoid a non-satisfying local minimizer. We applied the hard-thresholding algorithm (3.13) to the solution of the Forward-Backward algorithm (3.7) with various parameters  $\lambda_1, \lambda_2$  in order to get a number of nonzero frequencies smaller than the number of frequencies in tables 3.0(a) and 3.0(b), respectively six and two.

(a) Here we chose Lagrange parameters  $\lambda_1 = \lambda_2 = 3.10^3$  such that the numbers of nonzero frequencies in  $A_1$  and  $A_2$  are close to six as in table 3.0(a). The presence of outliers in the detection leads to significant errors in the least-squares solution.

Frequencies	Amplitudes ( $V_1$ )	Frequencies	Amplitudes ( $V_2$ )
0.006818	0.231948	0.000758	0.175933
0.007576	0.295975	0.003030	0.064870
0.009848	0.122875	0.004545	0.558571
0.015909	0.119733	0.008333	0.391177
0.020455	0.196495	0.018182	0.123960
		0.021212	0.258301
		0.021970	0.249609

(b) Elimination of outliers :  $\lambda_1 = \lambda_2 = 5.10^3$ . There are still outliers ( $\{0.006818, 0.000758\}$ ) in the detected frequencies but they have smaller amplitudes (0.093 and 0.187).

Frequencies	Amplitudes ( $V_1$ )	Frequencies	Amplitudes ( $V_2$ )
0.006818	0.093781	0.000758	0.187512
0.007576	0.347388	0.004545	0.501771

(c) Elimination of outliers :  $\lambda_1 = \lambda_2 = 9.10^3$ . Here we eliminated all the outliers and the associated amplitudes of the right frequencies match the data in table 3.0(a) (the phase opposition seems to be for 0.004545) but the correct frequencies are separated.

Frequencies	Amplitudes ( $V_1$ )	Frequencies	Amplitudes ( $V_2$ )
0.007576	0.343936	0.004545	0.531971

Table 3.2: Results of hard-thresholding iterations for the PLEIADES #1 images. The correct frequencies are 0.0048672 and 0.008112 with amplitude 0.25, we observe here that these two modes are correctly estimated. The first mode is 0.004545 with amplitude  $\approx 0.5$ , this seems to correspond to the phase opposition frequency, and the second mode is detected with a 0.007576 value and 0.36 amplitude instead of two components with 0.25 amplitude. It clearly appears that we recovered  $V_1 + V_2$  here instead of  $V_1$  and  $V_2$ .

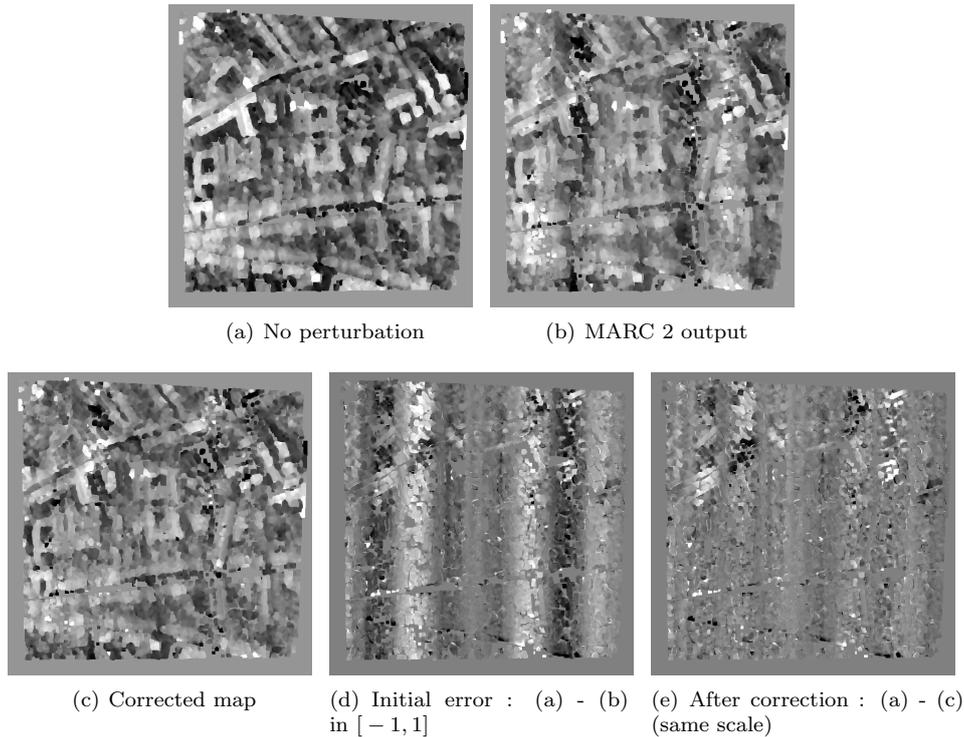


Figure 3.20: Correction of PLEIADES #1 map with parameters  $\lambda_1 = \lambda_2 = 9.10^4$  (see table 3.1(c)). Initial error oscillates between  $[-1, 1]$  (except for a few points due to the median filter). After correction the vertical stripes disappeared and the error oscillates between  $[-0.2, 0.2]$ .

Correction with parameter  $\lambda_1 = \lambda_2 = 9.10^3$  (no outlier) of PLEIADES images #1 is displayed in figure (3.20). Although we cannot quantify the accuracy of the correction, we observe that the vertical stripes have disappeared (we applied our correction to the median map for display purpose).

Experiments in table 3.2 and 3.3 show the importance of correct detection in the resolution of the Basis Pursuit problem. Amplitudes associated to outliers are significant in the PLEIADES images when the number of nonzero coefficients correspond to the modelization of microvibrations. It is necessary to over-regularize the microvibrations in order to eliminate these outliers and even in this case the problem of mixing between  $V_1$  and  $V_2$  can be confusing. Table 3.2 shows that an intermediary regularization factor ( $5.10^4$ ) reduces the  $l^0$  values to 2 which is the number of principal frequencies in  $V_1$  and  $V_2$ , but these frequencies are not present in both vectors. Other coefficients correspond to outliers (with small amplitudes). Table 3.2(b) reveals that without the stationarity hypothesis, the detected frequencies may contain outliers even if the number of frequencies matches the input data. Unfortunately we observed that adding this hypothesis does not eliminate the outlier 0.006838 as shown in table 3.4.

(a) Hard-thresholding with  $\lambda_1 = \lambda_2 = 2.10^4$ . The  $l^0$  value of the solutions is increased by the hard-thresholding step compared to the minimizer of  $\mathcal{P}_1$  (figure 3.15) and outliers have a great amplitude.

Frequencies	Amplitudes ( $V_1$ )
0.004883	0.763256
0.007812	0.390295
0.009766	0.657341

Frequencies	Amplitudes ( $V_2$ )
0.006836	0.995422
0.011719	0.585008

(b) Elimination of outliers :  $\lambda_1 = \lambda_2 = 3.10^4$ . Although the  $l^0$  values after hard-thresholding iterations are smaller than the correct values (table 3.0(b)) , there is still an outlier with significant amplitude (0.006836) which dominates the correct frequency 0.007812.

Frequencies	Amplitudes ( $V_1$ )
0.004883	0.903166
0.007812	0.339373

Frequencies	Amplitudes ( $V_2$ )
0.006836	0.734917

Table 3.3: Hard-thresholding for the Village #1 couple

Frequencies	Amplitudes ( $V_1$ )
0.004883	0.30
0.006836	0.34

Table 3.4: Hard-thresholding with stationarity hypothesis and same thresholds as in (3.2(b)). The solution does not improve the correction of the disparity map (see figure 3.16).

As a conclusion of this estimation step, we observed that in the provided PLEIADES images the  $l^1$  detection followed by a hard-thresholding step generally gives correct results provided that the number of frequencies is reduced to the number of principal modes (here 0.0048 and 0.0081). Frequencies with lower amplitudes are not detected in the first step and cannot be retrieved by the  $l^0$ -constrained least squares. Even in this case, the absence of outliers is not guaranteed because of the mixing between  $V_1$  and  $V_2$  but this phenomenon can be highly reduced by over-regularization. In that case we tried to eliminate outliers thanks to an *a contrario* approach, in the framework of *a contrario* methods proposed by Desolneux et al. in the late 90's.

### 3.3.2 A contrario least-squares

The *a contrario* method developed by Desolneux, Moisan and Morel [36] - [37] aims at detecting meaningful events (with a very low probability of occurrence) in images among criteria of the Gestalt theory of Wertheimer (1923) et al. . The main idea is that some qualitative geometric properties (alignment, parallelism, grouping, symmetry ...) are meaningful for the human vision system because they have a low probability of occurrence in the general set of possible images. In this context, measuring the meaningfulness of a geometric event requires an a priori statistical model for the general set of images. The classical example [36] is the detection of a square in an image whose pixels are the realization of i.i.d. Bernoulli variables  $\mathcal{B}(1, p)$ . Although this approach was originally applied to geometric problems, it has been successfully applied more recently to very different problems (for example [82] [31]) with the same formalism (the choice of a statistical model and a definition of  $\epsilon$ -meaningfulness). Work in this section is greatly inspired by works of L. Moisan and S. Durand in the one dimensional case [80], which deals with randomly sampled signals perturbed by additive Gaussian white noise. In that case it revealed quite efficient in recovering sums of few sine functions, as predicted by the Compressive Sensing theory.

Assume ( $H_0$ ) that the disparity  $D_k$  is the realization of i.i.d. Gaussian variables with mean zero and standard deviation  $\sigma$  (the mean part of the disparity is subtracted in the algorithms in order not to detect the zero frequency). For a given set of  $p$  frequencies  $F = \{f_j\}_{j=1..p} \subset ]0, \frac{1}{2}[$ , the minimization of the least squares problem :

$$I_F = \min_{\text{supp}(A), \text{supp}(B) \subset F_p} \|L(A, B) - D\|_2^2 \quad (3.16)$$

gives the squared norm of the difference between  $D$  (real) and its projection on the  $4p$   $\mathbb{R}$ -dimensional space spanned by :

$$\left\{ \left( e^{i2\pi f_j k_x} \right)_{k \in \Lambda}, \left( e^{-i2\pi f_j k_x} \right)_{k \in \Lambda}, \left( e^{i2\pi f_j k'_x} \right)_{k' \in \Lambda'}, \left( e^{-i2\pi f_j k'_x} \right)_{k' \in \Lambda'} \right\}_{j=1..p}$$

Thus the random variable  $\frac{1}{\sigma^2} (\|D\|_2^2 - I_F)$  follows a  $\chi^2$  law with  $4p$  parameters. In the spirit of a contrario methods, a quantity called *NFA* (*Number of False Alarms*) is associated to such a set  $F$  with  $|F| = p$  and represents a kind of probability of occurrence (the smaller the NFA, the less probable the event).

$$NFA(F) = N_{test}(p) \left( 1 - \beta_{4p} \left( \frac{1}{\sigma^2} (\|D\|_2^2 - I_F) \right) \right) \quad (3.17)$$

where  $N_{test}$  is the number of tests for objects similar to  $F$  (here the number of frequency subsets with  $p$  elements) and  $\beta_{4p}$  is the repartition function of the  $\chi^2$  law with  $4p$  parameters. Consider a maximal set of possible frequencies  $F_{max}$  with  $N_{max}$  elements, the total number of subsets with  $p$  elements is simply  $\binom{N_{max}}{p}$ . The function  $1 - \beta_{4p}(x)$  has the expression :

$$\frac{1}{(2p-1)!} \int_{x/2}^{+\infty} e^{-t} t^{2p-1} dt$$

and its logarithm can be approximated for large  $x$  by

$$\log_{10}(1 - \beta_{4p}(x)) \approx -(2p-1) \log_{10} \left( \frac{x}{2} \right) + \frac{x}{2 \ln(10)} + \log_{10}((2p-1)!)$$

which gives a fast computation of  $\log_{10}(NFA(F))$ .

An initial maximum set  $F_{max}$  is selected from the solutions of the detection step (with  $\ell^1$  regularization) and all subsets of  $F_{max}$  are tested. The computation of the  $NFA$  for each subset requires the resolution of a least squares problem with very few unknowns. In our problem, the number of selected frequencies is usually small ( $\leq 10$ ) and it is possible to compare the NFAs of all subsets of  $F_{max}$  in a reasonable time. We applied this algorithm to PLEIADES images with an initial set of frequencies reduced to approximately ten frequencies (table 3.2 (top) is a good example of the set of frequencies we used). The resolution of the least squares problem (3.16) in this case produces microvibrations  $V_1$  and  $V_2$  completely different from the results of the hard-thresholding algorithm. Here the vectors we obtain have high amplitudes and similar phases (table 3.5).

(a) A contrario method applied to image PLEIADES #1. Here the best subset found is the whole initial set of frequencies and amplitudes are completely overestimated by the least squares solving.

Frequencies	$ A_1(l) $	$\arg A_1(l)(rad)$	$ A_2(l) $	$\arg A_2(l)(rad)$
0.00681818	30.0695	2.82185	30.3029	2.80227
0.000757576	252.913	-0.0261188	252.959	-0.0283931
0.0030303	16.4557	-1.68534	16.511	-1.69452
0.00454545	12.0895	-1.77027	12.1493	-1.78369
0.00833333	37.4141	-0.368789	37.8258	-0.393235
0.0212121	2.96079	-2.44681	3.0136	-2.50594
0.0219697	1.69744	-0.853755	1.74718	-0.925229
0.00757576	53.9689	-1.90772	54.4271	-1.92992
0.00984848	17.272	-0.418362	17.6003	-0.447587
0.0106061	10.1413	1.11991	10.3687	1.08782
0.0204545	2.06628	2.34208	2.06584	2.28599

(b) With a different initial frequency set (left column) the overestimation of amplitudes persists.

Frequencies	$ A_1(l) $	$\arg A_1(l)(rad)$	$ A_2(l) $	$\arg A_2(l)(rad)$
0.00757576	5.53161	-1.58311	5.27379	-1.57022
0.00454545	4.44838	-2.89816	4.85415	-2.84339

Table 3.5: A contrario elimination of outliers : this approach fails because of the intrinsic formulation of our problem. The very special shape of the sampling set makes it highly ill-posed, the elevation  $D$  and the perturbations are not decorrelated.

In our experiments with the least squares method, the estimated microvibrations  $V_1$  and  $V_2$  were always overestimated and the residue  $D + V_1 - V_2$  was very small, which means that a large portion of the disparity can be represented with a few nonzero coefficients.

Indeed, problem  $(\mathcal{P}_1)$  admits trivial solutions : let  $V_1(x) = V_2(x) = -x + c^{te}$  on the disparity

support  $\Lambda \subset [0, n - 1] \times [0, m - 1]$ , then

$$V_1(z_k) - V_2(z_k - (D_k)_1) = -(D_k)_1 \quad \text{for all } 1 \leq k \leq |\Lambda|$$

Trivial solutions correspond to  $2T$  periodic functions equal to  $-x + c^{te}$  on the image domain. Since we usually take  $T = n$ , these trivial solutions are possible and unfortunately discredit the least squares method. This drawback cannot be eliminated by taking a different initial frequency set as shown in table 3.4(b) where we eliminated low frequency outliers (among which the largest harmonic components of the trivial solution).

This phenomenon can be explained by Parseval's energy formula : the solution of the least-squares problem with frequency set  $F$  has a great probability (depending on the lacunary set  $\Lambda$ ) to give frequency vectors  $(A_1, A_2)$  equal to the restriction on  $F$  of the trivial solution ( $V_1 = V_2 = -x + C^{te}$ ). This is not exactly the case in table 3.4(b), we remark that amplitudes are still over-estimated but lower than in table 3.4(a), we guess the shape of the sampling set  $\Lambda$  is the cause of this observation. The less probable frequency set in the a contrario approach will always be the whole initial set  $F_{max}$  which gives the smallest residu among all subsets of  $F_{max}$ . This explains why this approach does not work here. This also reveals that the separation of two microvibrations is much harder than the single microvibration estimation. In the Basis Pursuit and Hard-Thresholding steps, the presence of a non-quadratic regularization ( $\ell^1$  and  $\ell^0$ ) prevented this ill-posedness, the counterpart is the mixing phenomenon between  $V_1$  and  $V_2$  (a correct frequency lies in  $V_1$  or  $V_2$  but not in both  $V_1$  and  $V_2$  at the same time).

### 3.3.3 Conclusion of the estimation step

Our initial goal was the blind separation of the disparity and the microvibrations in the spirit of Basis Pursuit and our results show that this task is partially feasible in the framework of this study. Detection of microvibrations in PLEIADES type images is possible with this approach because the amplitudes of the disparity and the perturbation are not so different (a factor 10 is still reasonable) and more importantly because the contents of the disparity maps (buildings, streets, hills ...) do not correspond to the principal vibration modes. Estimation of amplitudes is more difficult and one of our approaches (Hard-Thresholding) gives encouraging results although we cannot measure the accuracy of the correction in the PLEIADES images. The a contrario approach revealed the high ill-posedness of the problem and the role of nonquadratic regularizations ( $\ell^1$  and  $\ell^0$ ) prevented this ill-posedness but the counterpart is the mixing between  $V_1$  and  $V_2$ .

We applied our blind separation method to synthetic data (images Village from the simulator) and observed that the resolution of the buildings in this case has an influence on the detection. The elevation model itself contains low frequencies because of the spacing between buildings and our corrections could not eliminate the outliers. Our observations also show that the separation of  $V_1$  and  $V_2$  in the images we considered is too difficult and that only  $V_1 - V_2$  can be estimated. In the next section we detail a successful attempt to deal with images of the Village type, in which large areas corresponding to affine elevation are present. It corresponds to the case when blind separation fails because of the elevation contents, but a closer look at the disparity map 3.3(c) reveals a segmentation of the disparity due to the adhesion correction mask. We propose a new formulation to exploit this feature.

### 3.4 An additional hypothesis for better correction

The preceding sections have shown the difficulty of estimating the amplitudes in the signals  $V_1$  and  $V_2$  from lacunary disparity maps. In the case of simulated images *Village*, this difficulty persists despite the presence of large affine parts in the map  $H(k+V_1(k))$  and this for several reasons. First, the disparity may contain low frequencies, like repetitive buildings structures or natural oscillatory elevations (sand dunes, hills, mountains). Second, the output of a sub-pixel disparity software like MARC2 is irregularly sampled, large holes may be present in the disparity mask which make the frequency detection harder. In this context of impossible blind separation, we an a priori on the elevation model  $H$  and consider it as a piecewise affine function in order to reduce the difficulty of this problem.

#### 3.4.1 Formulation of the piecewise affine hypothesis

Consider equation (3.6) :

$$D(z_k) = H(z_k + V_1(z_k)) - V_1(z_k) + V_2(z_k - D(z_k)) \quad 1 \leq k \leq |\Lambda|$$

The main difficulty in the previous modelizations is that the real disparity  $H$  may have large amplitudes compared to  $V_1$  and  $V_2$ , contains low frequencies as well as  $V_1$  and  $V_2$ , and that the sampling set  $\Lambda$  may have large gaps which remove frequency information. If  $H$  is locally determined by a few coefficients, the system (3.6) will have more equations (  $|\Lambda|$  ) than unknowns ( the coefficients of  $H$ ,  $A$  and  $B$  in a certain basis). Again, the sparsity of the different signals in presence is the key point we want to exploit to solve a problem apparently underdetermined.

This approach is somehow similar to the state of the art methods of section 3.1.2 in which a plane area is used to estimate the microvibrations, but it is more general. In urban areas, the disparity is certainly not a single plane on a length of several microvibrations periods, at least for the PLEIADES data. Assuming that the disparity  $H$  is piecewise affine seems more reasonable and the relevance of this assumption can be verified by looking at the elevation  $H(k+V_1(k))$  provided by the simulator (figure 3.21). We display in the middle the difference  $H(z_k) - H(z_k + V_1(z_k))$  thresholded between  $[-0.1, 0.1]$ . We verify that irregular sampling has a small influence on the shape of the elevation in flat areas and that most of the points in image 3.21(b) with high amplitudes correspond to edges of buildings and are eliminated by the adhesion correction filter (figure 3.21(c)).

The adhesion correction mask not only removes points with wrong disparities, but also separates the different surfaces present in the elevation (roofs, walls, ground). The remaining connected components generally correspond to a single surface and are very likely to be piecewise affine in urban areas. In dimension two, a function  $F$  is said to be affine if there exists  $a, b, c \in \mathbb{R}$  such that :

$$F(x, y) = a x + b y + c \quad (3.18)$$

and a differential characterization of such functions is :

$$F \in \mathcal{C}^2(\mathbb{R}^2) \quad \frac{\partial^2 F}{\partial x^2} = \frac{\partial^2 F}{\partial y^2} = \frac{\partial^2 F}{\partial x \partial y} = 0 \quad \text{on } \mathbb{R}^2 \quad (3.19)$$

At this point, our choice to treat the affine property in a variational formulation is questionable. Using either (3.18) on each connected component or (3.19) on the whole disparity domain would give the same function. We chose to use the variational characterization because of its flexibility. Adhesion correction here is a blessing since it removes edges in the disparity map associated to strong contrast changes in the original images, but it might not remove roof edges (discontinuity in the gradient of the elevation) and smooth non-affine surfaces, although this kind of surfaces

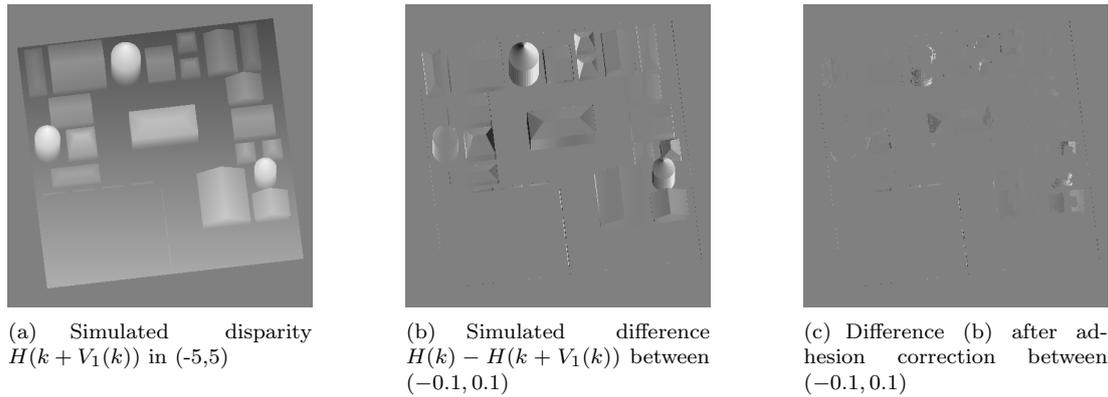


Figure 3.21: Relevance of the hypothesis on  $H(k + V_1(k))$ . The approximation  $H(k) \approx H(k + V_1(k))$  does not seem to introduce errors in the equation (3.6) thanks to adhesion correction which removes edges from the sampling set  $\Lambda_1 = \{k + V_1(k)\}$ .

is more unusual in satellite images (see for example the circular buildings from image 3.3(c) in the zoom below). The use of penalization on second order derivatives in (3.19) preserves some of the features that do not exactly fit the piecewise affine model, but the price for this flexibility is the number of unknowns in  $H$  (which is not reduced), the presence of one or more regularization parameters, and the necessity to solve a minimization problem involving second order derivatives on a lacunary sampling set.

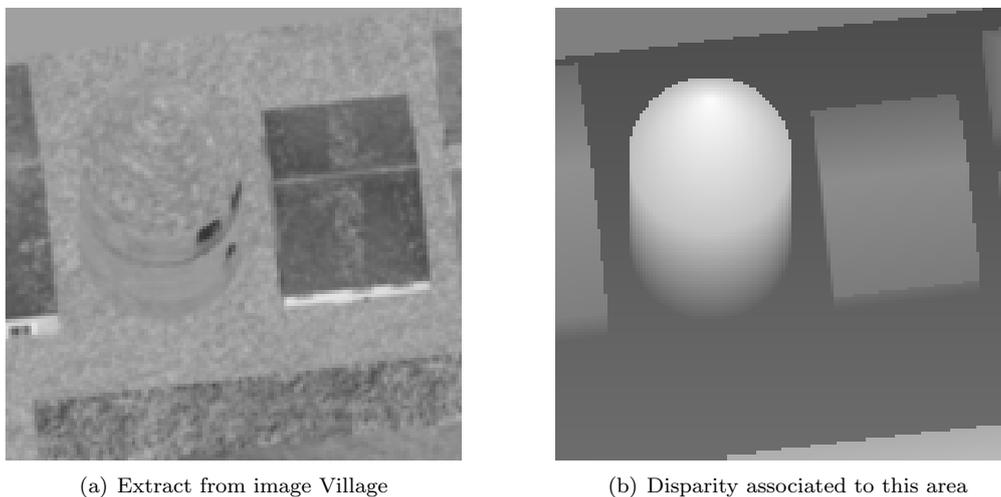


Figure 3.22: Example of nonaffine surfaces from images Village. Curved surfaces and roof edges do not satisfy the piecewise affine hypothesis and may be present in the final disparity map

### 3.4.2 Regularization on the second order derivatives

#### 3.4.2.1 Choice of an $\ell^1$ -norm

As mentioned earlier, our approach consists in decreasing the number of unknowns in our problem by approximating the disparity with a function close to (3.19). Unfortunately, penalizing the second order derivatives of  $H$  and assuming that  $H$  is piecewise affine are two very different ways of addressing the problem and not every penalization function will produce the desired result. The definition set of  $H$  is an important argument for using spatial penalization functions, as opposed to Fourier semi-norms. The regularity of the solution is the most important parameter, and we chose here to penalize the  $\ell^1$ -norm of the Hessian.

$$\sum_{k \in \Lambda_3} \sum_{\substack{0 \leq i, j \leq 2 \\ i + j = 2}} \left| \frac{\partial^2 H}{\partial x^i \partial y^j}(z) \right| \quad (3.20)$$

where  $\Lambda_3 \subset \Lambda$  is the set of points where these derivatives can be computed. In the theory of Sobolev spaces, this amounts to consider a solution in  $BV^2(\mathbb{R}^2)$  where  $BV^2(\Omega)$  is defined for regular open sets  $\Omega \subset \mathbb{R}^n$  by

$$BV^2(\Omega) = \left\{ u \in \mathcal{W}^{1,1}(\Omega) \text{ s.t. } \frac{\partial u}{\partial x_i} \in BV(\Omega) \quad \forall i \in \{1 \dots n\} \right\} \quad (3.21)$$

which allows discontinuous derivatives but does not allow discontinuous functions because of the following properties :

$$BV(\mathbb{R}^2) \not\subset \mathcal{C}(\mathbb{R}^2) \quad , \quad BV^2(\mathbb{R}^2) \subset \mathcal{W}^{1,2}(\Omega) \subset \mathcal{C}(\mathbb{R}^2)$$

where  $\mathcal{W}^{1,2}(\Omega)$  is the classical  $\mathcal{H}^1(\Omega)$  Hilbert space. (see [9]). This kind of Sobolev inclusions give a good insight of the properties of a differential semi-norm in penalization methods. In a discrete framework, these inclusions mean that discontinuities in the function have a much larger cost than discontinuities in the gradient. Another important property in the discrete implementation of  $\ell^1$  norms is the sparsity of the solution due to the non-smoothness of the  $\ell^1$  norm at points with a null coordinate. This sparsity is well-known in total variation regularizations (stair-casing effect in  $BV(\mathbb{R}^2)$  with the Total Variation) and in our case this is likely to produce solutions with sparse second order derivatives as desired.

#### 3.4.2.2 The Hessian operator and related variational problems

In the following we use the notation  $F$  for the piecewise affine disparity on  $\Lambda$ . We use the classical first and second order finite differences operators on images defined by

$$\left\{ \begin{array}{l} \frac{\partial^2 F}{\partial x^2}(k, l) = F(k-1, l) - 2F(k, l) + F(k+1, l) \\ \frac{\partial^2 F}{\partial y \partial x}(k, l) = \frac{\partial^2 F}{\partial x \partial y}(k, l) = F(k, l) - F(k-1, l) - F(k, l-1) + F(k-1, l-1) \\ \frac{\partial^2 F}{\partial y^2}(k, l) = F(k, l-1) - 2F(k, l) + F(k, l+1) \end{array} \right. \quad (3.22)$$

and note  $\mathcal{H}(F)$  the discrete Hessian matrix of  $F$ . Here the definition set of  $F$  is  $\Lambda \subsetneq [0, n-1] \times [0, m-1] \cap \mathbb{N}$  on the contrary of classical denoising problem and computation of the previous differences is not always possible. The penalization of the Hessian will involve a different set  $\Lambda_{sub} \subset \Lambda$  defined by :

$$\Lambda_{sub} = \{(k, l) \in \Lambda \times [0, n-1] \text{ s.t. } (k-1, l-1), (k, l), (k, l+1), (k-1, l), (k+1, l), (k-1, l-1) \in \Lambda\}$$

This corresponds to an erosion of  $\Lambda$  by the set of points  $\mathcal{N} = \{(-1, -1), (-1, 0), (0, -1), (0, 0), (0, 1), (1, 0)\}$  (see example in figure 3.23). Thanks to the practical observations on  $H$  (figure 3.21(c)) we can approximate  $F$  on the sampling set  $\Lambda_1 = \{k + V_1(k)\}$  as in equation (3.6) by its values on  $\Lambda$  :

$$F(k + V_1(k)) \approx F(k) \quad k \in \Lambda$$

We denote by  $M$  the linear operator from  $R^{|\Lambda|} \rightarrow \mathbb{R}^{3|\Lambda_3|}$  of second order finite differences

$$M F = \begin{pmatrix} \frac{\partial^2 F}{\partial x^2} \\ 2 \frac{\partial^2 F}{\partial y \partial x} \\ \frac{\partial^2 F}{\partial y^2} \end{pmatrix}$$

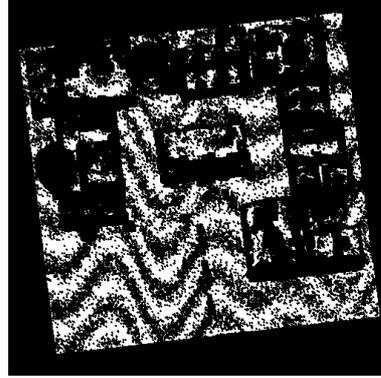
and we formulate our problem as a constrained minimization problem as in  $(\mathcal{P}_i)_{i=1,2,3}$

$$\begin{aligned} (A_1, A_2, F) &= \underset{(A,B,G)}{\operatorname{argmin}} E'(A, B, G) \text{ s.t. } \|A\|_1 \leq C_1, \|B\|_1 \leq C_2 \text{ and } \|\mathcal{H}(F)\|_1 \leq C_4 \\ \Rightarrow \exists \lambda_1, \lambda_2, \lambda_4 > 0 \text{ s.t. } (A_1, A_2, F) &= \underset{(A,B,G)}{\operatorname{argmin}} E'(A, B, G) + \lambda_1 \|A\|_1 + \lambda_2 \|B\|_1 + \lambda_4 \|\mathcal{H}(F)\|_1 \end{aligned} \quad (\mathcal{P}_4)$$

with  $E'(A, B, G) = \|D + L_1 A_1 - L_2 A_2 - G\|_2^2$  the squared differences on  $\Lambda$  and  $\|\mathcal{H}(F)\|_1$  the sum of absolute values of the Hessian on  $\Lambda_3$ . We made this approximation in the first place to simplify the problem in the  $F$  variable, but a clever implementation of a sampling operator on  $F$  would preserve the complexity of the problem (for example an interpolation operator with  $P_1$  finite elements, or a simple neighbourhood weighting). But the dependence of  $\Lambda_1$  on  $V_1$  would also change the properties of the minimized function in  $(\mathcal{P}_4)$ , in particular the convexity in the variables  $A, B$  and  $F$ .



(a) Original mask of points in  $\Lambda$



(b) Mask of points in  $\Lambda_3$  obtained by erosion

Figure 3.23: Set of sampling  $\Lambda$  of  $F$  (left) and set of penalization  $\Lambda_3$  on the Hessian  $\mathcal{H}(F)$  (right)

As in the detection section, it is possible to consider time dependant microvibrations as in  $(\mathcal{P}_3)$ ,

this leads to a simpler problem with only one vector of Fourier coefficients  $A_1$

$$\boxed{(A_1, F) = \underset{(A, G)}{\operatorname{argmin}} E'(A, S_\tau A, G) + (\lambda_1 + \lambda_2) \|A\|_1 + \lambda_4 \|\mathcal{H}(F)\|_1} \quad (\mathcal{P}_5)$$

The energies we deal with still have good minimization properties, we show in the next section that these new problems share many properties with previous problems  $(\mathcal{P}_i)_{i=1,2,3}$ , the same kind of tools can be used to solve them.

### 3.4.2.3 Minimization algorithms

We use the following notations :

$$E_4(A, B, G) = E'(A, B, G) + \lambda_1 \|A\|_1 + \lambda_2 \|B\|_1 + \lambda_4 \|\mathcal{H}(F)\|_1 \quad (3.23)$$

$$E_5(A, G) = E'(A, S_\tau A, G) + (\lambda_1 + \lambda_2) \|A\|_1 + \lambda_4 \|\mathcal{H}(F)\|_1 \quad (3.24)$$

Similarly to  $E_i, i \in \{1, 2, 3\}$ , the functions  $E_4$  and  $E_5$  belong to the nice class of  $l^1 + (l^2)^2$  functions (up to linear operators). Here we have :

$$E_4(A, B, G) = \|(\lambda_1 A, \lambda_2 B, \lambda_4 M G)\|_1 + \frac{1}{2} \|D + L_1 A - L_2 B - G\|_2^2 \quad (3.25)$$

with different vector sizes in the  $l^1$  and  $(l^2)^2$  terms, indeed  $A$  and  $B$  have size  $T$ ,  $M G$  the vector of second order derivatives has length  $3|\Lambda_3|$  and  $D$  has length  $|\Lambda|$ . Semi-implicit schemes and their accelerated versions are of course efficient algorithms here, but were not the first ones we considered. Experiments proved we were right when using an alternate minimization scheme for this problem (see algorithm 3.27). Still, it suffers from an intrinsic drawback of iterated dual schemes (we detail this later in the alternate minimization paragraph). Both the semi-implicit algorithm 3.29 and the alternate minimization algorithm share a common step which is the computation of the proximal of  $\|M \cdot\|_1$

$$\operatorname{prox}_{\lambda \|M \cdot\|_1}(Y) = \underset{G}{\operatorname{argmin}} \lambda \|M G\|_1 + \frac{1}{2} \|G - Y\|_2^2 \quad (3.26)$$

also called TV2-denoising problem (with quadratic fitting term). We first give the details of this minimization problem.

### 3.4.3 The TV2-denoising problem

This denoising problem has been addressed in the literature (Chambolle et al. 1997 [23], Piffet et al. 2010 [9]) for square images and is very close to our proximal problem in its implementation. The function

$$E_{TV2}(G) = \lambda \|M G\|_1 + \frac{1}{2} \|G - Y\|_2^2$$

is strictly convex on  $\mathbb{R}^{|\Lambda|}$  due to the square norm, but it is not differentiable because of the  $\ell^1$  term. Gradient and subgradient schemes have in this case a low convergence speed and the dual formulation of this problem provides faster algorithms (as in [22] with the Total Variation). The problem 3.26 can be formulated as a saddle point problem :

$$(\mathcal{P}_{TV2}) \quad \min_G \quad \max_{s \in K^*} \quad \lambda \langle M F, s \rangle_{\Lambda_3} + \frac{1}{2} \|Y - F\|_{\Lambda}^2 \quad (3.27)$$

and has the following dual problem (Fenchel-Moreau duality formula) :

$$(\mathcal{D}_{TV2}) \quad - \min_{s \in K^*} \frac{1}{2} \|Y - \lambda M^* s\|_{\Lambda}^2 - \frac{1}{2} \|Y\|_{\Lambda}^2 \quad (3.28)$$

Here  $K^*$  is the convex set of  $s \in (\mathbb{R}^3)^{|\Lambda_3|}$  such that  $|s_{i,j}| \leq 1 \quad i \in \{1, 2, 3\}, 1 \leq j \leq |\Lambda_3|$ . The vector  $F$  is expressed in terms of  $Y$  and  $s$  by  $F = Y - \lambda M^* s$ , that is the difference between  $Y$  and its projection on  $\lambda M^*(K^*)$ . Uniqueness of the minimizer in the dual problem is not guaranteed because of the definition of  $M^*$ , nevertheless the projection of  $Y$  on the convex set  $\lambda M^* K^*$  exists and is unique (theorem 9).

The algorithms we used to solve this dual problem are direct implementations of recent works in convex optimization by Combettes et al. [30], Nesterov [69], Beck et al. [8]. We describe below some Projected-Gradient-like algorithms (PG, FISTA) and Nesterov's algorithm applied to our problem.

0. Initialize  $n = 0, t_0 = 1, w^0 = s^0 \in (\mathbb{R}^3)^{|\Lambda_3|}, F^0 = \lambda (Z - M^* s^0), Z = \frac{1}{\lambda} Y$  and  $l > 0$  the Lipschitz constant of  $M M^*$  and  $0 < \rho \leq \frac{1}{l}$ .

1. **Repeat**

- (a)  $s^{n+1/2} = w^n - \rho M (M^* w^n - Z)$
- (b)  $s^{n+1} = \text{proj}_{K^*}(s^{n+1/2})$
- (c)  $t_{n+1} = \frac{1}{2} (1 + \sqrt{1 + 4t_n^2})$
- (d)  $w^{n+1} = s^{n+1} + \frac{t_n - 1}{t_{n+1}} (s^{n+1} - s^n)$
- (e)  $F^{n+1} = \lambda (Z - M^* s^{n+1})$
- (f)  $n \leftarrow n + 1$

**while**  $|E_{TV2}(F^n) - E_{TV2}(F^{n-1})| > \epsilon E_{TV2}(F^{n-1})$  and  $n < n_{max}$

2.  $G = F^n = Y - \lambda M^* s^n$

Figure 3.24: FISTA algorithm (accelerated version of the classical Projected Gradient algorithm here) applied to problem  $(\mathcal{D}_{TV2})$ . The stopping criterion is based here on the primal energy  $E_{TV2}$ .

Compared to [9] our utilization of the  $BV^2$  semi-norm is very different. First, the data has support in a lacunary subset of  $\Omega = [0, m - 1] \times [0, n - 1] \cap \mathbb{N}$ . But the goal of the regularization is also different, here we seek a strong regularization effect on the contrary of image denoising where the regularization is usually small (for reasonable noise). Convergence speed of the FISTA and Nesterov algorithms are equivalent ( $O(\frac{1}{N^2})$  with  $N$  the number of iterations) but Nesterov algorithm has a bigger computational cost per iteration. Indeed, it contains two Projected Gradient steps and twice the number of vectors as FISTA.

We display in figure (3.26) a comparison of Projected Gradient, FISTA and Nesterov algorithms applied to our problem with a small regularization parameter (non-monotonicity phenomena also occur with large regularization parameters). It is a crucial disadvantage of accelerated FISTA algorithm (and possibly Nesterov's algorithm, although it was not observed in this case). As mentioned by their authors (Beck [7] and Nesterov [70]) these algorithms have indeed a  $O(\frac{1}{N^2})$  convergence speed but do not necessarily decrease the primal energy in (3.26). Two monotone algorithms can be derived from FISTA and Nesterov's algorithms, as described in chapter 2 p. 31 and we display the energy decrease of these two monotone versions in figure (3.26). Monotone FISTA (mFISTA) generally converges faster than Nesterov's algorithm and has a lower computation cost per iteration, this justifies our preference for mFISTA in the following.

```

0. Initialize  $n = 0$ ,  $x^0, z^0, s^0 \in (\mathbb{R}^3)^{|\Lambda_3|}$ ,  $G^0 = 0 \in (\mathbb{R}^3)^{|\Lambda_3|}$ ,  $Z = \frac{1}{\lambda}Y$ ,  $F^0 = \lambda(Z - M^* s^0)$  and  $0 < \rho \leq \frac{1}{l}$  with  $l > 0$  the Lipschitz constant of  $MM^*$ .
1. Repeat
  (a)  $g^n = M(M^* x^n - Z)$ 
  (b)  $s^{n+1} = \text{proj}_{K^*}(x^n - \rho g^n)$ 
  (c)  $F^{n+1} = \lambda(Z - M^* s^{n+1})$ 
  (d)  $G^{n+1} = G^n + \frac{n+1}{2} g^n$ 
  (e)  $z^{n+1} = \text{proj}_{K^*}(x^0 - \frac{1}{l} G^{n+1})$ 
  (f)  $x^{n+1} = \frac{2}{n+3} z^n + \frac{n+1}{n+3} s^{n+1}$ 
  (g)  $n \leftarrow n + 1$ 
  while  $|E_{TV2}(F^n) - E_{TV2}(F^{n-1})| > \epsilon E_{TV2}(F^{n-1})$  and  $n < n_{max}$ 
2.  $G = F^n = Y - \lambda M^* s^n$ 

```

Figure 3.25: Nesterov's algorithm applied to problem  $(\mathcal{D}_{TV2})$ . It is very different from the FISTA algorithm as it uses a summation of all gradients from previous steps and an additional projection involving the initial vector.

It is commonly admitted among the Image Processing community that a denoising or deblurring algorithm does not need exact convergence, which justifies the use of relatively slow algorithms (sub-linear convergence speeds like  $O(\frac{1}{N^\alpha})$  with  $\alpha \geq 1$  are commonly used). The most important criterion is usually the numerical complexity of each iteration as a function of the number of unknowns. In our case the TV2 denoising problem is intended to be used iteratively, this is why we need a fast convergence rate.

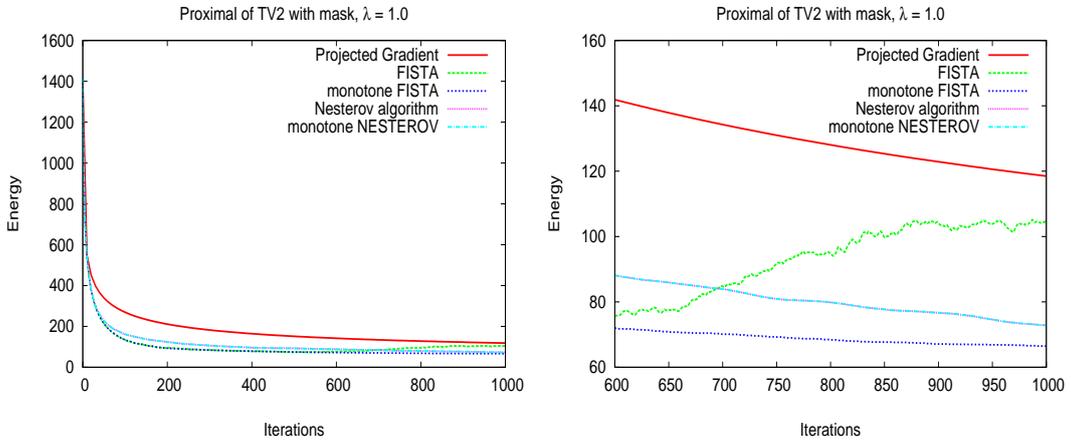


Figure 3.26: Energy in the primal problem  $\mathcal{P}_{TV2}$  with different methods. The non-monotonicity of FISTA algorithm at large iterations can completely cancel the gain at early iterations compared to Projected Gradient or Nesterov algorithms.

As described in [7], the monotone FISTA algorithm is guaranteed to produce a decreasing sequence  $(E_{TV2}(F_k))_{k \geq 0}$ . A stopping criterion can be adapted in the case of energy equality

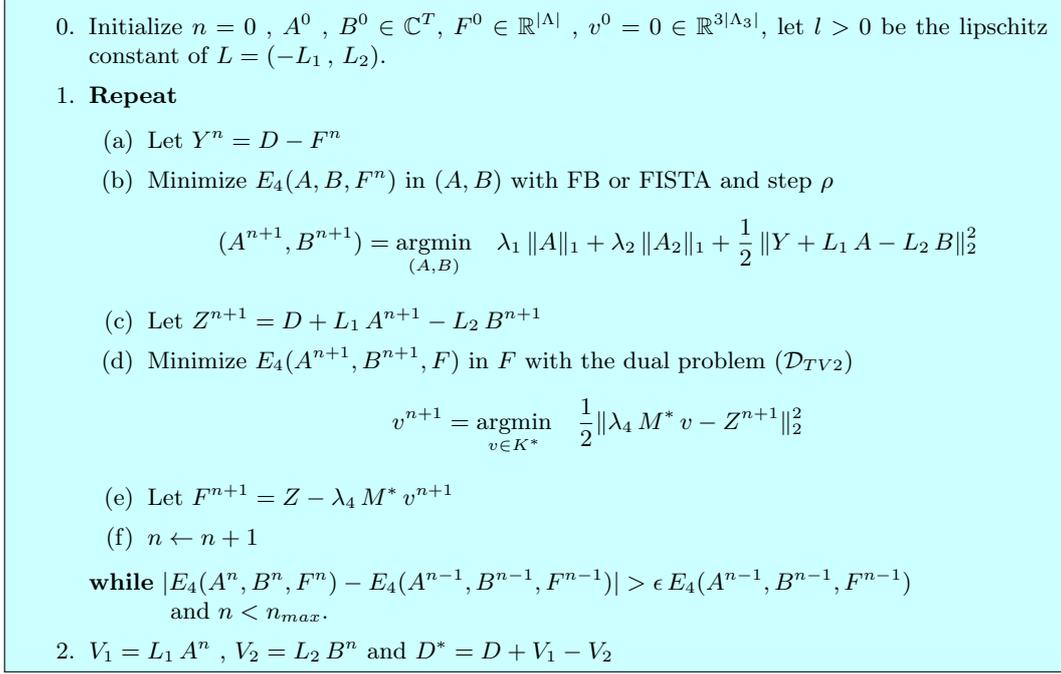
$E_{TV2}(F_{k+1}) = E_{TV2}(F_k)$  but the number of iterations necessary to decrease the energy can be arbitrarily large as we observed in some experiments with  $l^1$  minimization of problem  $(\mathcal{P}_1)$ . Nevertheless, monotone FISTA seems to be the best alternative to the minimization of  $E_{TV2}$  as it reaches a numerical precision much higher than Projected Gradient and Nesterov algorithms at early iterations.

### 3.4.4 Global minimization

#### 3.4.4.1 Alternate minimization

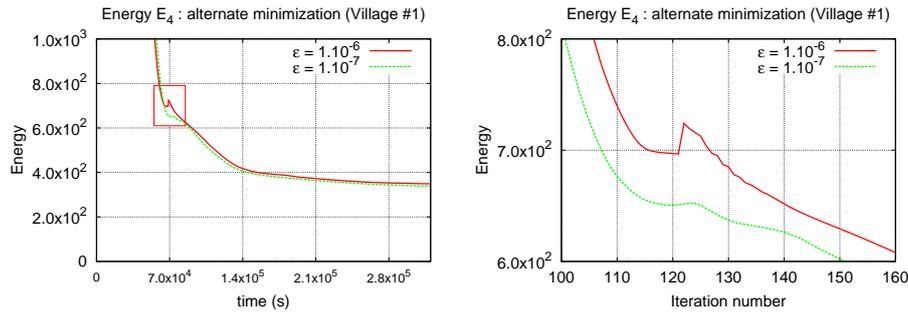
We first present the alternate minimization as it is the first algorithm we used and the one used to obtain the presented results, but a global scheme has also been tested, with surprising slower convergence. The scheme of the algorithm is presented in algorithm 3.27. The TV2 problem here is solved iteratively with a large parameter  $\lambda_4$  and it is applied to the vector  $Z^n = D + L_1 A_1^n - L_2 A_2^n$ . We present in figure 3.28 an example of energy decrease in the outer loop (the alternate minimization) at the origin of a very long debugging step.

In the dual problems we were faced with, the primal variable is always expressed as a function of the dual variable (the one we iteratively compute) for example  $U^* = V - \lambda \nabla^* s^*$  in the TV- $\ell^2$  problem. We observed in this experiment that the initialization of the dual variable can increase the energy  $E_4$ , a very large number of iterations (here  $10^5$  iterations) in the TV2 problem is the necessary in order to actually decrease the primal energy  $E_4$ .

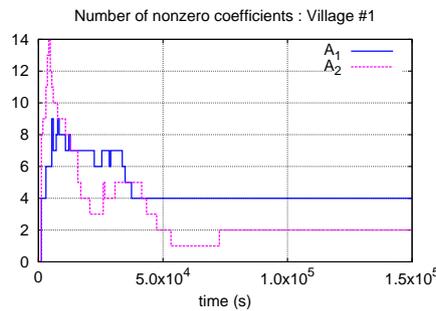


- Let  $\tilde{v}^0 = v^n$
  - $\tilde{F}^0 = Z^{n+1} - \lambda_4 M^* \tilde{v}^0$
  - Bad things can happen :  $E_4(A^n, B^n, \tilde{F}^0) > E_4(A^n, B^n, F^n)$

Figure 3.27: Alternate minimization of problem ( $\mathcal{P}_4$ ) (up). Iterated dual schemes (for the TV2 problem here) suffer from an initialization drawback since the primal energy is not preserved after  $F^n$  is updated. The number of iterations necessary to decrease the primal energy becomes arbitrarily large for large regularization parameters  $\lambda_4$ .



(a) Energy in the alternate minimization as a function of time (left) and zoom (right). Energy can increase due to the initialization in the dual inner loop under loose convergence requirements (although a  $10^{-7}$  relative convergence parameter is usually admitted).



(b) Evolution of the number of nonzero coefficients.

Figure 3.28: Energy in the alternate minimization scheme (left) and evolution of the number of detected frequencies (right) for  $\lambda_1 = \lambda_2 = 200$ . Despite the large number of iterations in the proximal it can happen that the proximal step increases the primal energy (here at approximately  $7.10^4$  seconds).

This could discredit the alternate minimization scheme, although we managed to get convergence with a very large number of iterations in the proximal step. A better understanding of the convergence criterion in the inner proximal loop would certainly prevent this huge computation cost, and works of Solodov et al. [86] for example on approximated proximal operators could be investigated.

### 3.4.4.2 Direct minimization schemes

The energies  $E_4$  and  $E_5$  belong to the same class of nonsmooth+smooth functions for which we used accelerated semi-implicit schemes. Algorithm 3.29 presents the classical Forward-Backward algorithm applied to this problem, and we refer to algorithm 2.5 p. 31 for the scheme of the accelerated version (this is valid for all Forward-backward algorithms presented earlier).

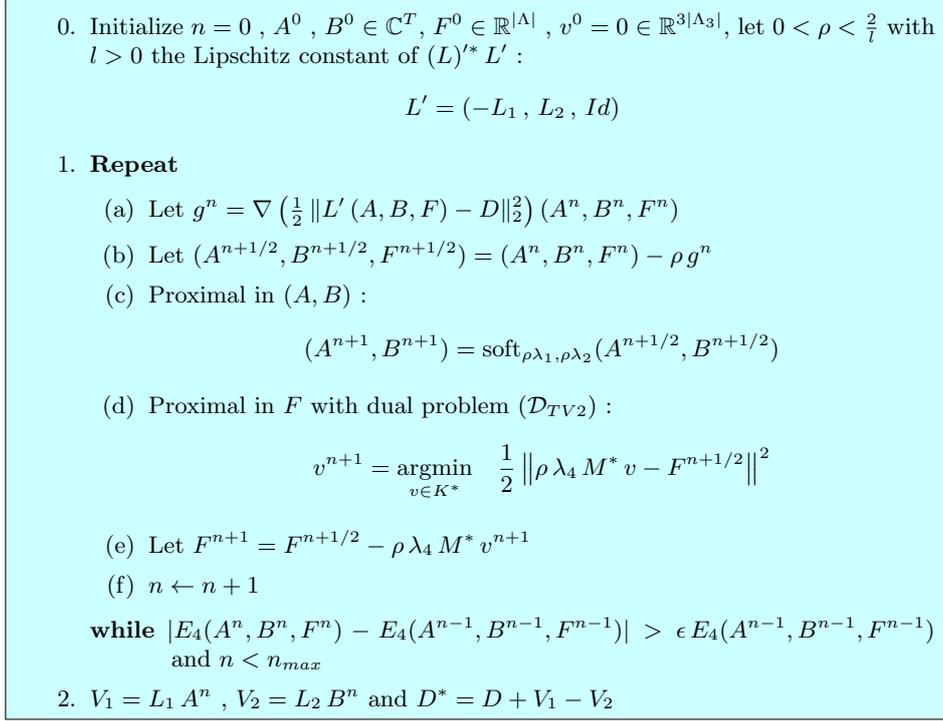


Figure 3.29: Direct minimization of problem  $(\mathcal{P}_4)$  with the FB algorithm (take  $\rho \leq \frac{1}{l}$  for the FISTA algorithm).

This algorithm contains the same kind of dual inner loop as the alternate minimization scheme but the proximal parameter (the factor  $\rho \lambda_4$ ) is much smaller. In our tests the order of magnitude of  $\lambda_4$  was  $1.10^4$  and  $\frac{1}{\rho} \in \{2.10^5, 4.10^5\}$ , the stopping criterion in the proximal was  $\epsilon = 1.10^{-6}$  and the maximum number of iterations  $n_{max} = 10^5$ .

## 3.4.4.3 Results

We display in figure (3.30) the corrected disparity maps and nonzero coefficients of vectors  $(A_1, A_2)$  obtained by the alternate minimization scheme on  $E_4$ . This shows that the piecewise affine hypothesis was correct in this case, the regularization parameter  $\lambda_4$  on the Hessian has to be quite large in order to separate the low-frequency microvibrations and the elevation part.

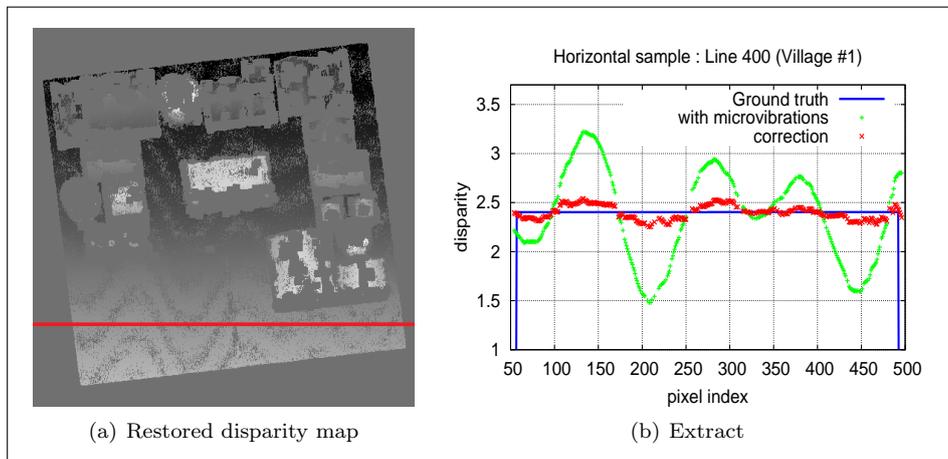


Figure 3.30: Experiment on the disparity map with adhesion correction (left) : We observe the separation of the affine elevation and the oscillating component. The small remaining oscillations (middle) are certainly due to the erosion effect of the  $l^1$  norm, but it outperforms the previous partial correction of figure 3.16 with much smaller Lagrange parameters  $\lambda_1$  and  $\lambda_2$ . ( $\lambda_1 = \lambda_2 = 200$ ). Microvibrations have been reduced by a factor 10.

The algorithm does not perform so well when there is no adhesion correction (figure 3.32). The regularization parameter  $\lambda_4$  has to be chosen smaller than previous experiment (figure 3.30), such parameters no longer work with the subpixel disparity map. Over-regularization in this case has the same drawback as previous detection and estimation methods, the elevation is not well fitted by the piecewise affine term  $F$  and a part of the elevation (mainly edges) will be fitted by the microvibration terms  $L_1 A_1$  and  $L_2 A_2$  instead.

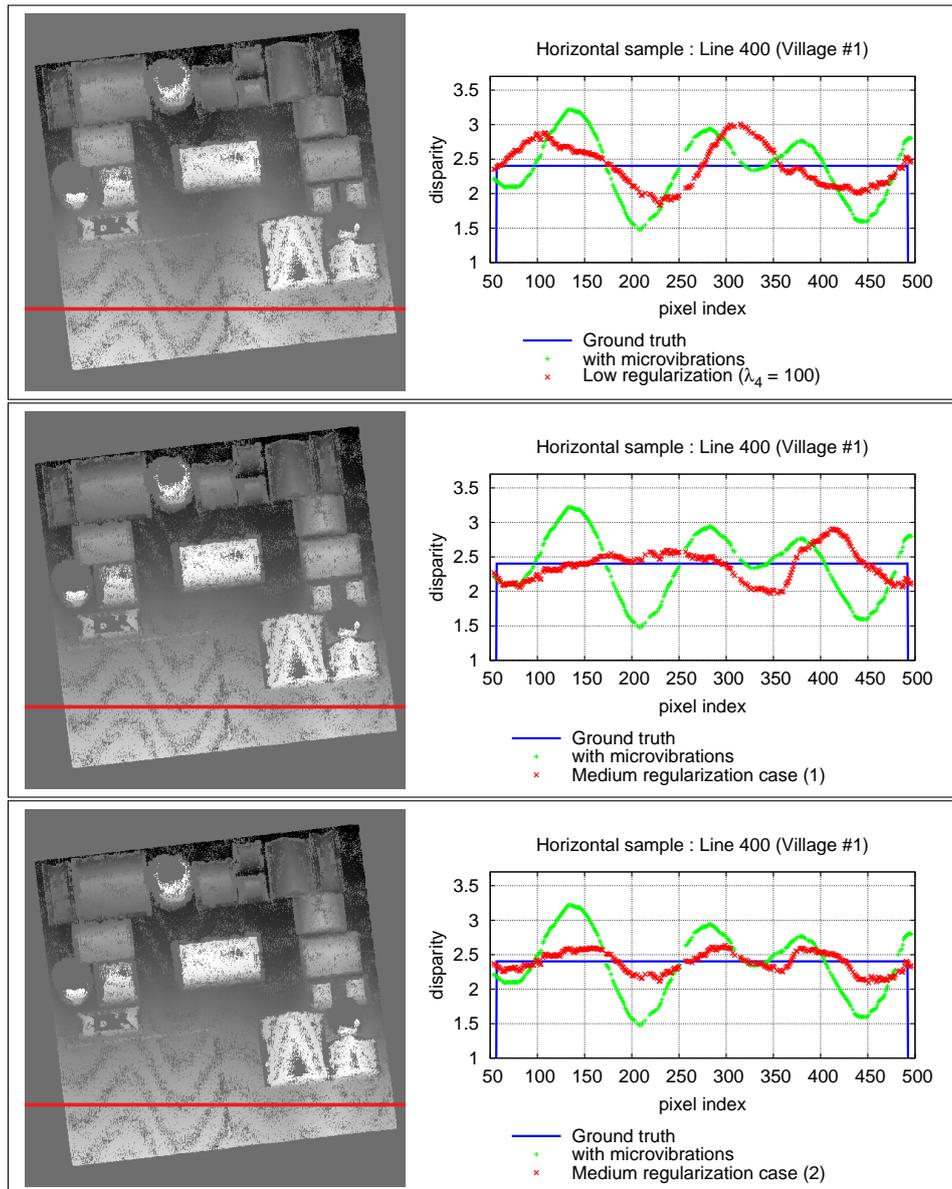


Figure 3.31: Experiments on the subpixel disparity map (without adhesion correction). For different regularization parameters  $\lambda_4 \in \{10^2, 10^3\}$  and microvibrations with  $\ell^1$  norm approximately equal to 1. The estimated disparity  $D + V_1 - V_2$  is clearly affected the discontinuities in the elevation due to buildings. Low regularization corresponds to  $\lambda_4 = 1.10^2$ ,  $\lambda_1 = \lambda_2 = 200$ , we observe that the lowest frequency  $0.0048 \text{ pix}^{-1}$  is not correctly detected in  $V_1$  and  $V_2$  but is present in  $F$ . Medium regularization case (1) and case (2) correspond to  $\lambda_4 = 1.10^3$ ,  $\lambda_1 = \lambda_2 \in \{400, 600\}$ , the frequency  $0.0048 \text{ pix}^{-1}$  is better detected.

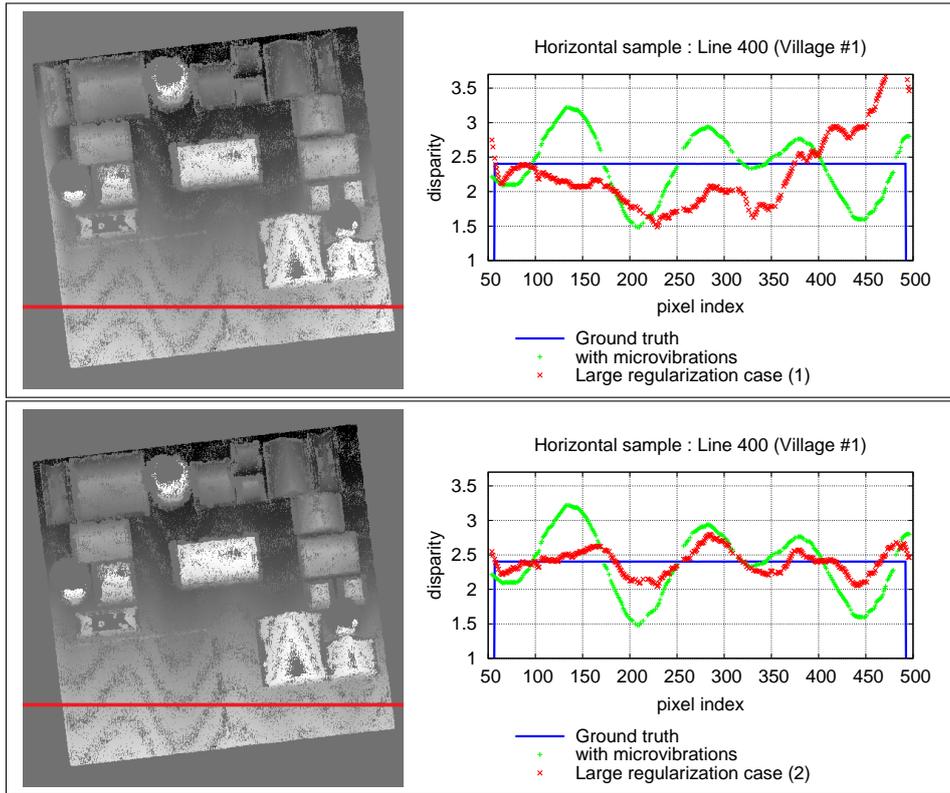


Figure 3.32: Experiments on the subpixel disparity map (without adhesion correction) with large regularization parameter  $\lambda_4 = 1.10^4$ . Cases (1) and (2) correspond to  $\lambda_1 = \lambda_2 \in \{2.10^2, 1.10^3\}$ . Case (1) has the same parameters as in figure 3.30, we clearly observe the bad behaviour due to elevation fitted by the microvibrations. Case (2) is better since we used a larger  $\ell^1$ -penalization parameter, amplitudes are closer to what they should be :  $\|A\|_1 + \|B\|_1 \approx 1$ .

#### 3.4.4.4 Numerical observations

The reason why we displayed results from the alternate minimization above is that the Forward-Backward and monotone FISTA in this case exhibited very bad convergence rates, the Forward-Backward was not even monotonous for all the parameters we tested. Figure 3.33 presents the energy convergence as a function of time for different convergence parameters  $\epsilon$  in the inner dual TV2 algorithm.

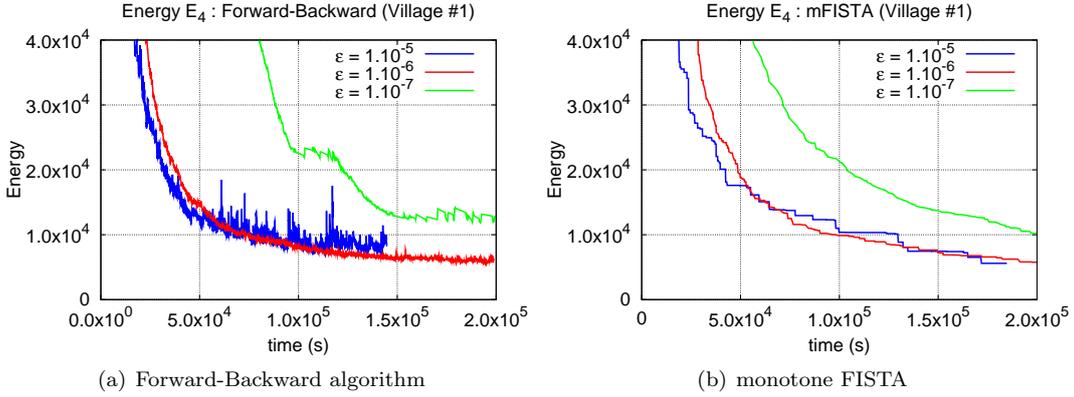


Figure 3.33: Influence of the convergence parameter in the dual inner loop for both Forward-Backward (FB) and monotone FISTA algorithms. FB is not monotonous but the chaotic behaviour is reduced with lower parameter  $\epsilon$  (with the drawback of an increased computation time). mFISTA can be non-decreasing during large periods of time with low convergence rate  $\epsilon = 1.10^{-5}$  but still performs better than cases  $\epsilon = 1.10^{-6}$  and  $\epsilon = 1.10^{-7}$  which are nearly monotonously decreasing.

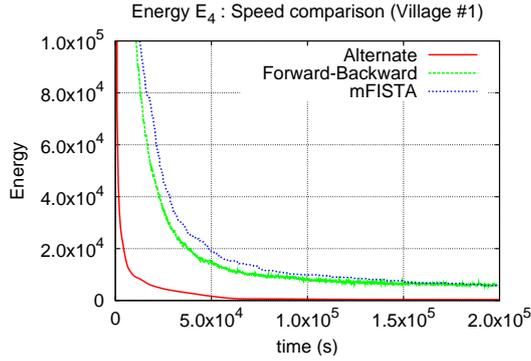


Figure 3.34: Energy for the three algorithms we implemented. Alternate minimization has much better performances in this case than our implementation of global schemes FB and mFISTA.

Results from the FB and mFISTA schemes were very different from the Alternate scheme because of the lack of convergence. The norm of the Hessian was indeed decreased and solutions looked piecewise affine, but the fitting constraint did not converge. This produces discontinuities between the different connected parts. Convergence of the TV2 prox in both FB and mFISTA does not seem to be crucial here (figure 3.33) and a better strategy could be to stop this inner loop after very few iterations.

But the weakness of the global minimization scheme here seems to be the large Lipschitz constant of  $L' = (-L_1 l_2 Id)$  compared to  $Id$  due to  $L_1$  and  $L_2$ . A change of variable  $F \leftarrow \frac{1}{C} F$  in this case amounts to minimize :

$$\operatorname{argmin}_{A,B,\tilde{F}} \quad \lambda_1 \|A\|_1 + \lambda_2 \|B\|_1 + \lambda_4 C \|M \tilde{F}\|_1 + \frac{1}{2} \|D + L_1 A - L_2 B - C \tilde{F}\|^2 \quad (3.29)$$

with  $C > 0$  the maximal Lipschitz constant of  $L_1$  and  $L_2$ .

We examined this formulation as a possible improvement of algorithm 3.29, we believe this would accelerate the algorithm as the formulation is more general than the alternate minimization scheme and involves almost the same Lagrange parameters :  $\lambda_4$  is replaced by  $\frac{\lambda_4 C}{l}$  with  $l > 0$  the Lipschitz constant of  $(-L_1 \ L_2 \ C Id)^* (-L_1 \ L_2 \ C Id)$ . This strategy belongs to the more general set of preconditioning methods, we believe this would make the Forward-Backward and FISTA more efficient than the alternate minimization scheme although here this necessarily modifies the Lagrange parameters and thus the convergence rate of the different loops.

### 3.5 Conclusion

In this study we treated the highly ill-posed problem of microvibrations in a couple of stereo images by means of non-quadratic regularization ( $\ell^1$  and then  $\ell^0$ ). Experiments show that it is indeed absolutely necessary because of the unusual form of the sampling set and the relation between this last one and the data. In a first detection step Basis Pursuit provided the two main frequency modes in the perturbation in the case of simulated PLEIADES images.

Over-smoothing can eliminate outliers but of course this is a difficult task since the frequency set is not known in general. Furthermore the mixing phenomenon between microvibrations  $V_1$  and  $V_2$  makes harder this elimination with a constraint based on the number of nonzero coefficients. We obtained a quite good estimation of the perturbation with a  $\ell^0$  constrained variational method (Iterated Hard-thresholding Algorithm) which is nonconvex and thus highly depends on the initialization and the presence of outliers.

The Village image was more challenging because of the correlation between microvibrations and elevation. Detection step could not get rid of some outliers this is why we proposed a new method with a penalization on the Hessian of the elevation. This worked very well when the adhesion mask gives a nice segmentation of the data, it is a little less efficient without this adhesion mask and raises a new numerical challenge, indeed the computation time was much larger than the single ISTA and HSTA algorithms.

A more general method for subpixel maps, which could be investigated together with numerical considerations, is the TV-TV2 inf-convolution from Chambolle et al. [23] which amounts to decompose the disparity into two terms plus the microvibrations :

$$D = V_2 - V_1 + F_1 + F_2$$

with  $F_1 \in BV(\Omega)$  and  $F_2 \in BV^2(\Omega)$ .

## Chapter 4

# Restauration of irregularly sampled images with splines

### 4.1 Introduction

This chapter deals with the very first topic treated during my thesis which pulled me to the satellite imaging domain. Among the classical problems in satellite imaging (calibration, denoising, deblurring, images fusion . . .) we find a more specific problem due to the push-broom acquisition mode of many satellites. For example the latest SPOT satellites launched by the CNES, and the new generation of PLEIADES satellites, use a TDI (Time Delayed Integrator) sensor with much more columns than rows (about ten rows and several thousands of columns). This sensor receives light from Earth locations which may not be exactly equally distributed because of the sensor's motion (trajectory) and microvibrations (unstable pointing direction). The resulting image is not sampled on the usual rectangular grid  $\Lambda_{reg} = \{(j, k) \in [0, m[ \times [0, n[ \cap \mathbb{N}^2\}$  but rather on a perturbation of this grid  $\Lambda$ .

$$\Lambda = \{(x + \epsilon_1(x, y), y + \epsilon_2(x, y)) \in \mathbb{R}^2 \text{ s.t. } (x, y) \in \Lambda_{reg}\}$$

with  $\epsilon = (\epsilon_1, \epsilon_2)$  a two dimensional perturbation. Here we consider more general two dimensional perturbations, on the contrary of previous section where we treated only the problem of pitch correction ( $\epsilon_2 = 0$ ). These perturbations have a low amplitude (a few pixels) and are low frequency (about a tenth of the sampling frequency of the image) but do not have a sparse Fourier transform. We study the general case of random perturbations  $(\epsilon_1, \epsilon_2)$  obtained by taking the real part of a two dimensional signal with random i.i.d. Gaussian Fourier coefficients  $\approx \mathcal{N}(0, 1)$  in the basis of complex exponentials.

We address here the inverse problem of recovering an image  $u$  irregularly sampled in the context of satellite imaging. We first describe the two problems (resampling and deblurring+resampling) and the literature in this domain. This literature is rich and already provides accurate algorithms based on the bandlimitedness hypothesis and local regularization. But these approaches suffer from the heavy computational cost of trigonometric polynomials in dimension two. We propose a resolution of this problem based on a different approximation space (splines). The first part of this work concerns the irregular sampling problem, this is a first approach treated during my training course. The second part includes deblurring with the Total Variation and a fast algorithm, part of this work was presented at the ICIP 2010 conference in Hong-Kong [3].

### 4.1.1 Two irregular sampling problems

In a very general image formation framework, we address both resampling ( $\mathcal{P}_{samp}$ ) and deblurring+resampling ( $\mathcal{P}_{blur}$ ) problems.

$$\boxed{\text{Find } u^* \text{ s.t. } u^*(z_k) = v_k + \nu_k \quad \text{for all } z_k \in \Lambda} \quad (\mathcal{P}_{samp})$$

$$\boxed{\text{Find } u^* \text{ s.t. } (h * u^*)(z_k) = v_k + \nu_k \quad \text{for all } z_k \in \Lambda} \quad (\mathcal{P}_{blur})$$

with  $u^* : \Omega = [0, n] \times [0, m] \rightarrow \mathbb{R}$ ,  $\{v_k\}_{k=1..|\Lambda|}$  the data,  $h$  a convolution kernel in either  $L^1(\mathbb{R}^2)$  or  $L^1(\Omega)$  depending on the boundary conditions and  $\nu_k$  the noise term (here i-i-d. Gaussian variables).

The first problem applies to a very wide range of images as the formation model is not taken into account. The second problem is more particular and here we consider here the satellite image formation model of SPOT 5 HRG hipermode described in [61].

$$h_{sat} = h_{opt} * h_{sens} * h_{mov} \quad (4.1)$$

namely the optical, sensor and motion equivalent PSFs with

$$\hat{h}_{opt}(\xi_1, \xi_2) = e^{-\alpha_1 \sqrt{\xi_1^2 + \xi_2^2}} \quad (4.2)$$

$$\hat{h}_{sens}(\xi_1, \xi_2) = \text{sinc}\left(\frac{\xi_1}{2}\right) \text{sinc}\left(\frac{\xi_2}{2}\right) e^{-\beta_1 |\xi_1|} \quad (4.3)$$

$$\hat{h}_{mov}(\xi_1, \xi_2) = \text{sinc}\left(\frac{\xi_1}{4}\right) \text{sinc}\left(\frac{\xi_2}{4}\right) \quad (4.4)$$

In particular we decompose the sensor PSF in two terms :

$$\hat{h}_{sens,1}(\xi_1, \xi_2) = \text{sinc}\left(\frac{\xi_1}{2}\right) \text{sinc}\left(\frac{\xi_2}{2}\right) \quad , \quad \hat{h}_{sens,2}(\xi_1, \xi_2) = e^{-\beta_1 |\xi_1|}$$

as  $\hat{h}_{sens,1}$  is the Fourier transform of a spline function and can be treated separately. The movement PSF also belongs to a spline space, indeed

$$h_{mov}(x, y) = \mathbb{1}_{[-1,1]}(x) \otimes \delta_0$$

and in the following we mainly focus on the non-spline PSF  $h = h_{opt} * h_{sens,2}$ . The TDI instrument also implies a spatially variant blurring kernel due to microvibrations of the array sensors  $h_{vib}$ , this kernel is unfortunately not shift-invariant and cannot be diagonalized in the Fourier basis this is why we did not consider the vibrations kernel in this work but could be integrated since we work with known perturbations  $\epsilon$  on the sampling grid.

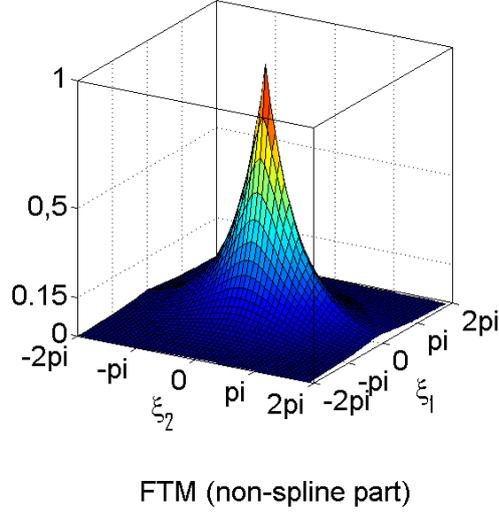


Figure 4.1: FTM of the convolution kernel  $h_{mov} * h_{opt}$  we consider in the following : it is at most 0.15 at the frequency  $(0, \pi)$  and approximately  $8 \cdot 10^{-2}$  at  $(\pi, 0)$ .

#### 4.1.2 State-of-the-art

Image resampling is not a recent problem, in fact it was already addressed in the early 90's by Feichtinger, Groechenig and Rauth [49] [87] and people have always searched for faster algorithms because of the large number of unknowns in images. In [50] the authors proposed an acceleration of the frame algorithm, in [87] the Toeplitz property of the sampling operator is used to solve the least-squares problem in the space of 2D trigonometric polynomials. In [48] the authors proposed an acceleration based on frame theory and local weights and named this method ACT (*Adaptive weights and Conjugate gradient with Toeplitz matrices*). The ACT algorithm also applies to certain quadratic regularizers if the Toeplitz structure is preserved. This is the case for the  $H^1$  or  $H^2$  discrete semi-norms, but more recently Almansa et al. have combined trigonometric polynomials and nonquadratic regularization to efficiently solve highly ill-posed problems (deblurring and zooming) [4] [46]. These algorithms are very accurate and stable, but their high computational cost is problematic, this is why we studied the resampling and deblurring+resampling problems in another functional setting (splines). We present below the ACT+TV (TV regularization with local constraints) algorithm in [4] and ACT+FAR (Frequency Adaptive Regularization with local constraints) from [46] the two state-of-the-art algorithms with nonquadratic regularization.

ACT+TV and ACT+FAR aim at recovering an image  $u$  from incomplete data

$$\{v_k = P(u)(\lambda_k)\}_{k=1..|\Lambda|}$$

on an irregular grid  $\Lambda = \{\lambda_k\}_{k=1..|\Lambda|} \subset \mathbb{R}^2$  where  $P$  is diagonal in the Fourier basis (convolution and/or spectral projection). The image is approximated by a two dimensional, real trigonometric polynomial :

$$u(x, y) = \sum_{k=(1-n)/2}^{(n-1)/2} \sum_{l=(1-m)/2}^{(m-1)/2} \hat{u}[k, l] e^{i(\frac{2k\pi}{n} + \frac{2l\pi}{m})}$$

with odd sizes  $n, m$ . The associated linear sampling operator is denoted by  $S_\Lambda$ , it acts on the Fourier coefficients  $\hat{u}[\cdot]$ , each of its blocks of size  $n \times |\Lambda|$  has Vandermonde structure up to a multiplication, and the overall product  $(S_\Lambda P)^*(S_\Lambda P)$  has a two dimensional Toeplitz structure (it is block-Toeplitz with Toeplitz blocks). The minimized energy in ACT+TV and ACT+FAR is the Lagrangian corresponding to the minimization of an energy  $E(u)$  under local constraints  $g_j(u) \leq 0$ . These local constraints tend to homogenize the residue  $(S_\Lambda P \hat{u} - v)$  locally to a value depending on the noise level.

$$g_j(x) = \left( G(k) * (S_\Lambda \hat{P} \hat{u} - v)_k^2 \right) [j] - \sigma^2 \quad , \quad j \in J$$

where  $G$  is a normalized, positive, narrow window function (indicator or Gaussian) and  $J$  a set of points homogeneously distributed in the image domain. The form of the minimized function is  $E(u) + \sum_{j \in J} \mu_j g_j(u)$  with  $\mu_j \geq 0$  the Lagrange multipliers. In [4] the function  $E(u)$  is the classical Total Variation of Rudin et al. [79], and in the ACT+FAR method  $E(u)$  is the discretization of a two-dimensional fractional Sobolev semi-norm.

$$E_{TV} = \sum_k \sum_l \|\nabla_{++} u(k, l)\| \quad (4.5)$$

$$E_{FAR} = \sum_k \sum_l \sqrt{\left\| H_\gamma(u) \left( \frac{k}{2}, \frac{l}{2} \right) \right\|^2 + C} \quad (4.6)$$

where  $\nabla_{++}$  is the classical discretized gradient operator,  $C > 0$  is a small regularization parameter and  $H_\gamma(u)$  is defined in the Fourier domain by

$$\mathcal{F}(H_\gamma(u))(\xi_1, \xi_2) = \left( |\xi_1|^\gamma \frac{i\xi_1}{|\xi_1|} \hat{u}(\xi_1, \xi_2), |\xi_2|^\gamma \frac{i\xi_2}{|\xi_2|} \hat{u}(\xi_1, \xi_2) \right)$$

(the case  $\gamma = 1$  is a Riemann series involving the exact gradient of  $u$  computed in the Fourier domain)

Function  $E_{TV}$  is not differentiable because of the norm of the gradient, this is why in [4] Almansa, Haro, Caselles and Rougé used a dual formulation.  $E_{FAR}$  is differentiable and the algorithm in this case is a Quasi-Newton algorithm (iterated minimization of a second order approximation involving the Hessian of  $E_{FAR}$ ). The algorithm developed in ACT+TV contains two nested loops (plus Uzawa's loop). The outer loop is an implicit diffusion (Crandall-Liggett) with the subgradient of  $E_{ACT+TV}$  :

$$\frac{u_{n+1} - u_n}{\tau} \in -\partial E_{ACT+TV}(u_{n+1}) \quad (4.7)$$

such that cluster points  $u^*$  of  $(u_n)_{n \in \mathbb{N}}$  verify  $0 \in \partial E_{ACT+TV}(u^*)$ . The developed algorithm solves the dual form of this equation, this amounts to compute the proximal of  $(TV(Q^{-1/2} \cdot))^*$  with  $Q$  a positive definite operator :

$$b_n \in w_{n+1} + \partial TV(Q^{-1/2} w_{n+1}) \quad \text{with } w_{n+1} = Q^{1/2} u_{n+1}, b_n = u_n + \tau Q^{1/2} v \quad (4.8)$$

This proximal computation is the inner loop and is performed by Chambolle's algorithm [22]. Unfortunately it amounts to inverse operator  $Q$  by conjugate gradient at each iteration. We give the outline of the algorithm in [4] (Uzawa's loop is not included) :

We notice that ACT+TV algorithm always requires an invertible operator  $Q = (S_\Lambda P)^*(S_\Lambda P)$ . The algorithm is quite slow (approximately 30 minutes for a  $149 \times 149$  image) in the deblurring case although the simple resampling algorithm takes a few seconds, this is why the ACT+FAR method was later proposed by Facciolo, Almansa, Aujol and Caselles in [46].

0. Initialize  $u_0$
1. For all  $n \geq 0$  compute  $u_{n+1}$  solving (4.7)
  - (a) Compute  $b_n$ , initialize  $\tilde{s}_0, k = 0$
  - (b) For all  $k \geq 0$  compute  $\tilde{s}_{k+1}$  by Chambolle's algorithm
    - Compute  $q_k = Q^{-1} \nabla_{++}^* \tilde{s}_k$  by the Conjugate Gradient algorithm
    - Compute the gradient  $g_k = \nabla_{++} (q_k + Q^{-1} b_n)$
    - Update  $s_{k+1}$
    - $k \leftarrow k + 1$
  - (c)  $u_{n+1} = Q^{-1}(b_n + \nabla_{++}^* \tilde{s}_{k+1})$
  - (d)  $n \leftarrow n + 1$

Figure 4.2: ACT+TV algorithm.

This ACT+FAR method is the one we use in our comparative tests since it is much faster and efficient than ACT+TV. The Quasi-Newton algorithm used in ACT+FAR is a fixed point where the denominator of  $\nabla E_{FAR}$  is assumed constant and each iteration amounts to solve a linear system (the approach is more conventional as it uses only linear algebra).

0. Initialize  $u_0, n = 0$
1. For all  $n \geq 0$  compute  $u_{n+1}$  by the Conjugate gradient algorithm solving :

$$\frac{H_\gamma^* H_\beta u_{n+1}}{\sqrt{\|H_\gamma u_n\|^2 + C}} + \sum_j \nabla g_j(u_{n+1}) = 0$$

2.  $n \leftarrow n + 1$

Figure 4.3: ACT+FAR algorithm.

## 4.2 Restoration with splines

The resampling problem ( $\mathcal{P}_{samp}$ ) in perturbed sampling does not really suffer from the slow convergence observed in previous methods. Though, the following section provides an interesting introduction to the state of the art methods with splines. Our resolution of the deblurring+resampling problem can be found in section 4.2.3 p. 89.

### 4.2.1 Stable resampling

We consider the case of nonblurry images, acquired at irregular locations  $\Lambda = \{(x_{k,l}, y_{k,l})\}$  with  $(x_{k,l}, y_{k,l}) = (k, l) + (\epsilon_1, \epsilon_2)(k, l)$ , and suffering from additive Gaussian white noise  $w_{k,l} \sim \mathcal{N}(0, \sigma)$

$$I_{k,l} = u(x_{k,l}, y_{k,l}) + w_{k,l} \quad \text{for all } k, l \quad (4.9)$$

If we assume that  $u$  belongs to a spline space, the recovering of  $u$  from the data  $I_{k,l}$  can be highly ill-posed even in the absence of noise because of the local sampling density of  $\Lambda$ . Indeed, the relation between the coefficients of  $u$  in a spline basis and the data  $I_{k,l}$  is a linear system with a few nonzero coefficients on each row, as opposed to the interpolation with trigonometric polynomials or thin-plate splines. If the sampling grid has a hole around a point  $(k_0, l_0)$  of radius  $\frac{d+1}{2}$  with  $d$  the degree of the regular B-splines, then the linear system is under-determined since the sampling matrix  $S_\Lambda$  has a null column. We write the relation between vector  $(I_{k,l})$  and the spline coefficients of  $u$  as a linear system :

$$(I_{k,l}) = S_\Lambda (a_{k,l}) \quad (4.10)$$

with  $(S_\Lambda)_{(k,l),(q,p)} = \beta_{d,2}(x_{k,l}-q, y_{k,l}-p)$  for all  $0 \leq k, q < n$  and  $0 \leq l, p < m$  if the approximation space is  $S^d(n, m)$ .

To illustrate this ill-posedness, we display in figure 4.4 an example of irregularly sampled image restored by a simple least-squares applied to (4.10). Artifacts are clearly visible in several parts of the restored image, although the quantity  $\|S_\Lambda a - I\|^2$  (the square norm of the residue) tends to zero. These artefacts are amplified in the presence of noise, this justifies the use of stable methods such as regularization. In the next section we present a quadratic regularization with a semi-norm on the Hessian of  $u$  inspired by works of Arigovindan et al. [5].

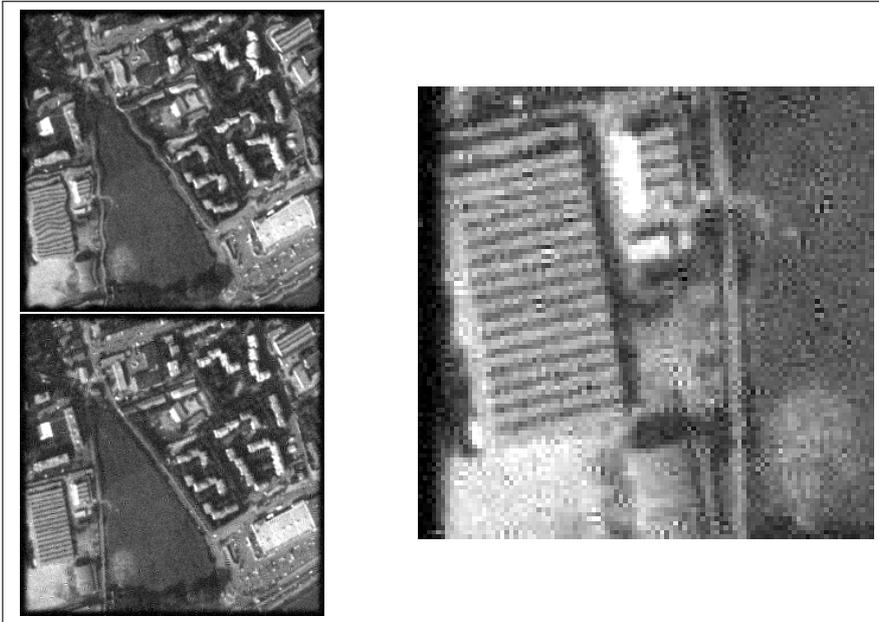


Figure 4.4: Distorted and noisy image (CNES copyrighted image) with noise level  $\sigma = 5$  (left) and solution obtained by least-squares solving of problem (4.10) (zoomed at right). We remark the presence of localized high-frequency artefacts due to the ill-conditioned sampling matrix  $S_\Lambda$ .

## 4.2.2 Quadratic regularization on the Hessian

### 4.2.2.1 Variational formulation

In article [5] Arigovindan, Sühling, Hunziker and Unser present a multigrid algorithm for solving the Lagrangian constrained problem (4.11), we give a few details of this implementation only for

the last scale. The formulation amounts to find  $u^*$  such that

$$u^* = \operatorname{argmin}_{v \in S^3(n,m)} \|v(x_{k,l}, y_{k,l}) - I_{k,l}\|^2 + \lambda \|D^2 v\|_{L^2(\mathbb{R}^2)}^2 \quad (4.11)$$

with an approximation space  $S^3(n, m)$  equal to the space of bicubic splines. This problem is well defined in the space  $S^3$  since bicubic splines are  $\mathcal{C}^2$  on  $\mathbb{R}^2$  and the minimizer is unique since constant and affine functions do not belong to  $S^3(n, m)$  whose elements are compactly supported in  $\mathbb{R}^2$ . The function  $E_H : v \rightarrow \|D^2 v\|$  is a thus norm on  $S^3(n, m)$  and  $E_H^2$  rewrites as a strictly convex quadratic function  $E_H(v) = \langle H a, a \rangle$  where  $a$  is the vector of spline coefficients of  $v$  and the positive definite matrix  $H$  can be computed by means of scalar products between shifted splines  $\beta_{d,2}$  and their derivatives :

$$H_{(k,l),(q,p)} = H_{(k,l),(q,p)}^{xx} + H_{(k,l),(q,p)}^{yy} + 2H_{(k,l),(q,p)}^{xy}$$

with

$$\begin{aligned} H_{(k,l),(q,p)}^{xx} &= \iint_{\mathbb{R}^2} \frac{\partial^2 \beta_{d,2}}{\partial x^2}(\cdot - (k, l)) \frac{\partial^2 \beta_{d,2}}{\partial x^2}(\cdot - (q, p)) \\ H_{(k,l),(q,p)}^{yy} &= \iint_{\mathbb{R}^2} \frac{\partial^2 \beta_{d,2}}{\partial y^2}(\cdot - (k, l)) \frac{\partial^2 \beta_{d,2}}{\partial y^2}(\cdot - (q, p)) \\ H_{(k,l),(q,p)}^{xy} &= \iint_{\mathbb{R}^2} \frac{\partial^2 \beta_{d,2}}{\partial x \partial y}(\cdot - (k, l)) \frac{\partial^2 \beta_{d,2}}{\partial x \partial y}(\cdot - (q, p)) \end{aligned}$$

The approximation space  $S^3(n, m)$  here might not be well adapted near the image boundaries where this space has different approximation properties due to missing knots around boundary samples. In our case this fortunately did not appear thanks to the windowing applied to the data. This windowing was used in previous state-of-the-art articles [4] [46] to avoid periodization problems, here we used it to prevent the mentioned problem in  $S^3(n, m)$ .

If the minimization is performed on  $S_{per}^3(n, m)$  instead of  $S^3(n, m)$  and the Hessian is integrated on  $\Omega = [0, n[ \times [0, m[$ , the regularity term is no more strictly convex, constant functions lie in the kernel of the Hessian operator but the function still admits a unique minimizer. Given two minimizers  $u^*$  and  $v^*$  the difference  $(u^* - v^*)$  would lie in the kernel of  $H : v \rightarrow D^2 v$  which consists in constant functions (or equivalently functions with constant spline coefficients). But the function  $G : v \rightarrow \|v(x_{k,l}, y_{k,l}) - I_{k,l}\|^2$  is strictly convex on any affine space oriented in the constant direction. Thus the problem (4.12) below has a unique solution.

$$u_{per}^* = \operatorname{argmin}_{v \in S_{per}^3(n,m)} \|v(x_{k,l}, y_{k,l}) - I_{k,l}\|^2 + \lambda \iint_{\Omega} \|D^2 v(x, y)\|^2 dx dy \quad (4.12)$$

The matrix  $H$  becomes circulant (the matrix of a convolution operator) and we verify that its kernel is restricted to constant functions by the observation of its singular values. This approach consists in the penalization of the square norm of the Hessian, and thus gives a solution  $u^*$  in  $\mathcal{W}^{2,2}(\mathbb{R}^2)$ . This assumption implies that both  $u^*$  and  $\nabla u^*$  are continuous by Sobolev inclusions and the result  $u^*$  will be smooth (i.e. blurry). This kind of regularization belongs to the family of Tikhonov methods [68] which usually oversmooths edges in images. In the simple resampling and denoising framework the authors of [5] claim that this penalization is sufficient and that bicubic splines provide a nearly optimal solution in reference to works of Duchon [40]. Duchon showed that the solution of problem (4.11) in the distributions sense is a particular thin-plate spline :

$$u(X) = P(x, y) + \sum_{(x_{k,l}, y_{k,l}) \in \Lambda} \phi((x, y) - (x_{k,l}, y_{k,l}))$$

where  $P$  is a polynomial belonging to the kernel of the Hessian operator and  $\phi$  a radial function whose expression depends on dimension  $d$  :

$$\phi(\|X\|) = \begin{cases} \|X\|^{4-d} & \text{if } 4-d \text{ is odd} \\ \|X\|^{4-d} \log(\|X\|) & \text{else} \end{cases}$$

In dimension one, and in the regular sampling case, cubic splines provide the optimal solution since  $x \rightarrow |x|^3$  coincides with a function in  $S^3(n)$  on  $[0, n[$ . This is not the case in dimension two, nevertheless bicubic splines have been used in many problems to approximate images and are very efficient in this problem. The quantities  $H^{xx}, H^{yy}, H^{x,y}$  at index  $(k, l), (q, p)$  only depend on  $(k-q, l-p)$  and the computation of  $H$  is reduced to the computation of three filters  $h^{xx}, h^{yy}, h^{xy}$ . Derivatives of splines can be expressed as linear combinations of splines with lower degree (formula (2.19) p. 20 ) and using the scalar product formula (2.20) we get :

$$\begin{aligned} h^{xx}(k, l) &= ((\delta_{-2} - 4\delta_{-1} + 6\delta_0 - 4\delta_1 + \delta_2) \otimes \delta_0) * (\langle \beta_1(\cdot), \beta_1(\cdot - k) \rangle \otimes \langle \beta_3(\cdot), \beta_3(\cdot - l) \rangle) \\ &= ((\delta_{-2} - 4\delta_{-1} + 6\delta_0 - 4\delta_1 + \delta_2) * \beta_3[k]) \otimes \beta_7[l] \\ h^{xy}(k, l) &= ((\delta_{-1} - 2\delta_0 + \delta_1) \otimes (\delta_{-1} - 2\delta_0 + \delta_1)) * (\langle \beta_2(\cdot), \beta_2(\cdot - k) \rangle \otimes \langle \beta_2(\cdot), \beta_2(\cdot - l) \rangle) \\ &= ((\delta_{-1} - 2\delta_0 + \delta_1) * \beta_5[k]) \otimes ((\delta_{-1} - 2\delta_0 + \delta_1) * \beta_5[l]) \\ h^{yy}(k, l) &= h^{xx}(l, k) \end{aligned} \tag{4.13}$$

Finally,  $H$  is the matrix of the convolution with the following kernel

$$h = h^{xx} + h^{yy} + 2h^{xy} = \begin{bmatrix} 0.0002050 & 0.0073016 & 0.0411706 & 0.0693122 & 0.0411706 & 0.0073016 & 0.0002050 \\ 0.0073016 & 0.0800000 & 0.0142857 & -0.2031746 & 0.0142857 & 0.0800000 & 0.0073016 \\ 0.0411706 & 0.0142857 & -0.6776786 & -0.2555556 & -0.6776786 & 0.0142857 & 0.0411706 \\ 0.0693122 & -0.2031746 & -0.2555556 & 3.4455028 & -0.2555556 & -0.2031746 & 0.0693122 \\ 0.0411706 & 0.0142857 & -0.6776786 & -0.2555556 & -0.6776786 & 0.0142857 & 0.0411706 \\ 0.0073016 & 0.0800000 & 0.0142857 & -0.2031746 & 0.0142857 & 0.0800000 & 0.0073016 \\ 0.0002050 & 0.0073016 & 0.0411706 & 0.0693122 & 0.0411706 & 0.0073016 & 0.0002050 \end{bmatrix} \tag{4.14}$$

The minimization of (4.11) amounts to solve the Euler-Lagrange equation below since both terms are differentiable and  $H$  is self-adjoint.

$$2 S_{\Lambda}^t (S_{\Lambda} a - I) + 2\lambda(H^t + H) a = 0 \quad \Leftrightarrow \quad a = (S_{\Lambda}^t S_{\Lambda} + \lambda H)^{-1} S_{\Lambda}^t I \tag{4.15}$$

We considered two approaches for the choice of Lagrange parameters. The first one is based on works of Vazquez et al. [93] who proposed a different formulation of this problem by means of local Lagrange parameters  $\alpha = (\alpha_{k,l})$  and the solving of :

$$S_{\Lambda}^t (S_{\Lambda} a - I) + \lambda D_{\alpha} H a = 0 \tag{4.16}$$

where  $D_{\alpha}$  is the diagonal matrix with general term  $D_{(k,l),(k,l)} = \alpha_{k,l} \geq 0$  (the  $\alpha_{k,l}$  depend on the sampling grid) and  $\lambda$  is the regular Lagrange parameter.

The second approach is based on a denoising criterion. We consider the constrained denoising problem :

$$u^* = \underset{u \in S^3(n,m)}{\operatorname{argmin}} \|D^2 u\|^2 \quad \text{s.t.} \quad \|u(x_{k,l}, y_{k,l}) - I_{k,l}\|^2 \leq R |\Lambda| \sigma^2 \tag{4.17}$$

where  $R$  is a denoising parameter chosen in  $]0, 1]$ . This last problem can be addressed by the Uzawa algorithm with a classic Lagrangian (with  $\lambda$  a multiplier of the regularization term as above) and

we consider a band of width  $W |\Lambda| \sigma^2$  for valid solutions of the constrained problem :

$$(u^*, \lambda^*) \text{ is a valid solution if } \begin{cases} u^* = \underset{u \in S^3(n,m)}{\operatorname{argmin}} \|u(x_{k,l}, y_{k,l}) - I_{k,l}\|^2 + \lambda^* \|D^2 u\|^2 \\ \|u^*(x_{k,l}, y_{k,l}) - I_{k,l}\|^2 \in [(R - W) |\Lambda| \sigma^2, (R + W) |\Lambda| \sigma^2] \end{cases} \quad (4.18)$$

In the following we detail this second approach and the results obtained with the conjugate gradient algorithm.

#### 4.2.2.2 Conjugate gradient and Uzawa's algorithm

The linear system in (4.15) is of the form  $Y = \Phi X$  with  $\Phi$  a symmetric positive definite matrix. We even know that  $\Phi X - Y$  is the gradient of the function  $F(a) = \frac{1}{2} (\|S_\Lambda a - I\|^2 + \lambda \langle H a, a \rangle)$  at  $X$ . We used the following conjugate gradient algorithm 4.5 with an initial vector  $x_0$  obtained by the simple Least-Squares and a stopping criterion based on the noise level :

- Let  $\tilde{a}_0 = 0$
- Let  $\tilde{a}_n$  be result of the CG algorithm with  $\lambda = 0$  ( $\Phi = S_\Lambda^t S_\Lambda$ ) and stopping criterion

$$\|S_\Lambda \tilde{a}_n - I\|^2 \leq (R + W) |\Lambda| \sigma^2$$

0. Let  $n = 0$ ,  $a_0 \in \mathbb{R}^{n \times m}$ ,  $\Phi = S_\Lambda^t S_\Lambda + \lambda_k H$ ,  $r_0 = S_\Lambda^t I - \Phi a_0$ ,  $p_0 = r_0$
1. Compute  $a_{n+1}$  by the CG algorithm 2.2 with stopping criterion
 
$$(R - W) |\Lambda| \sigma^2 \leq \|S_\Lambda a_n - I\|^2 \leq (R + W) |\Lambda| \sigma^2$$

Figure 4.5: Conjugate gradient algorithm applied to problem (4.15) with Lagrange parameter  $\lambda_k$  ( $k^{th}$  step of the Uzawa algorithm).

The computation cost of this initialization is lower than the regularized CG since only  $S_\Lambda$  and the Hessian filter  $h$  are computed. In our observations it produces quite good initial images, but remaining distortions and noise make this approach convenient only as an initialization. We can derive an estimate of the Lagrange parameter  $\lambda^*$  from  $a_0$  by :

$$S_\Lambda^t (S_\Lambda a_0^* - I) + \lambda^* H a_0^* \approx 0 \quad \Rightarrow \quad \lambda^* \approx \lambda_0 = \frac{\|S_\Lambda^t (S_\Lambda a_0^* - I)\|}{\|H a_0^*\|} \quad (4.19)$$

and we chose to update the Lagrange parameter with an exponential method as in Tseng [88] :

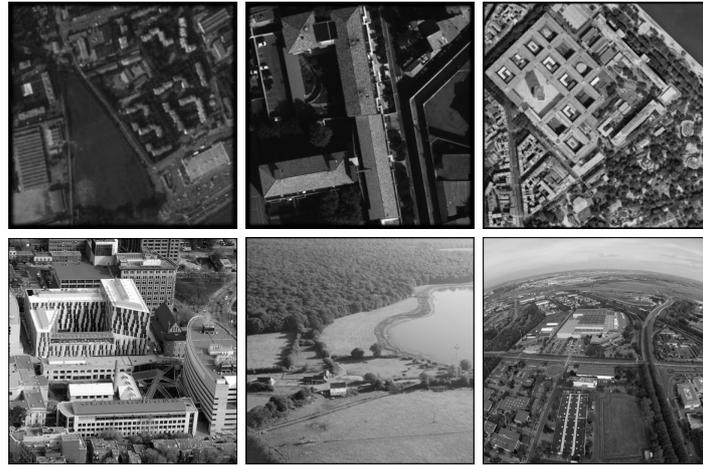
$$\lambda_{k+1} = \lambda_k / \min\{e^{\|S_\Lambda a_k^* - I\|^2 - R |\Lambda| \sigma^2}, c_0\} \quad (4.20)$$

with  $a_k^*$  the result of the CG algorithm at step  $k$ .

This update of the Lagrange parameter is in fact a semi-implicit version of the original Uzawa algorithm with exponential constraint described in [88] and produces the desired couple  $(u^*, \lambda^*)$  with a very small number of iterations in Uzawa's algorithm. We present some of the test images we used in figure (4.7) together with the SNR results and computation times. The value of  $R$  depends of course on the image but also on the noise level. For low noise levels a factor  $R \approx 0.5$  gives the best performances, but this factor increases for higher noise levels from 0.5 to approximately 0.8. This phenomenon is well known and is simply due to the variational method which cannot fully denoise a signal and preserve its contents, such that partial denoising is generally preferred.

0. Let  $\tilde{a}_0 = 0$ ,  $\mu = |\Lambda|\sigma^2$ ,  $k = 0$ .
1. Compute  $a_0^*$  by the CG algorithm 4.5 with  $\lambda = 0$  and stopping criterion
 
$$\epsilon = (R + W)\mu$$
2. Compute  $\lambda_0$  with equation (4.19)
3. Repeat
  - (a) Compute  $a_{k+1}^*$  by the CG algorithm with  $\lambda_k$  and stopping criterion
 
$$\epsilon \ll W\mu$$
  - (b) Update  $\lambda_k$  by formula (4.20)
  - (c)  $k \leftarrow k + 1$
4. While  $\|S_\Lambda a_k^* - I\|^2 \notin [(R - W)\mu, (R + W)\mu]$

Figure 4.6: Uzawa algorithm with exponential constraint and multiplicative update applied to the stable resampling problem.



(a) Images with various sizes and contents (CNES copyrighted images. (1-2))

image	W = 0.1			W = 0.01		
	time(s)	SNR	RMSE	time(s)	SNR	RMSE
nimes	0.680	28.90	1.09	1.212	28.94	1.08
stmichel	2.612	32.46	1.02	4.520	32.47	1.02
jussieu	2.024	31.43	1.43	3.620	31.50	1.42
uqam	10.417	31.02	1.73	53.471	31.07	1.73
lac	2.972	31.53	1.13	9.129	31.61	1.12
parisnord	5.216	32.31	1.38	5.216	32.31	1.38

(b) SNR results obtained with the quadratic Hessian regularization (top-left to bottom right above images)

Figure 4.7: Test images and results with a noise level  $\sigma = 1$  and a partial denoising criterion  $R = 0.5$ . A narrow band stopping criterion  $W$  increases the computation time with a negligible improvement in the RMSE. (Single thread implementation run on a dual-core 2.5 GHz CPU)

image	$\sigma = 3$			$\sigma = 5$			$\sigma = 10$		
	time(s)	SNR	RMSE	time(s)	SNR	RMSE	time(s)	SNR	RMSE
nimes	0.520	22.21	2.35	0.504	19.5	3.38	0.476	14.57	5.67
stmichel	2.104	25.70	2.23	1.976	22.43	3.25	1.832	17.79	5.55
jussieu	1.344	25.81	2.74	1.132	22.55	3.99	1.080	18.38	6.53
uqam	5.240	26.74	2.84	7.584	23.68	4.04	4.144	19.39	6.63
lac	2.228	25.34	2.31	1.948	22.18	3.33	1.856	17.52	5.69
parisnord	3.924	26.85	2.58	3.280	23.76	3.68	3.092	19.40	6.09

Figure 4.8: Results of the Hessian regularization for  $R = 0.5$  and  $W = 0.1$  and noise levels  $\sigma \in \{3, 5, 10\}$ . Computation time depends on the noise level : partial denoising not only gives good SNR results but also accelerates the restoration.

This first approach of the sampling problem showed that the computational cost of a stable spline approximation was reasonable (it is a  $O(m \times n)$  with  $n, m$  the image sizes) thanks to the fast conjugate gradient algorithm and a strongly convex regularization. But quadratic regularizers are known to be very bad at denoising very noisy signals because of the smoothing effect. In that case a nonlinear regularization is necessary, in the following we apply the TV regularization to  $(\mathcal{P}_{blur})$ .

### 4.2.3 On the deblurring problem

Works of Almansa et al. [4] [46] have shown the improvements in image quality of nonquadratic regularizers such as TV or TV-FAR with local constraints. These methods benefit from the setting of bandlimited functions, in particular convolution operators are well defined and have a convenient expression in the Fourier domain. This is not the case with splines and we begin this section with a presentation of the new mathematical formulation of the deblurring problem  $(\mathcal{P}_{blur})$  in spline spaces.

#### 4.2.3.1 Convolution and projection on spline space

Spline functions  $\beta_{d,2}$  are not bandlimited functions, indeed the Fourier transform

$$\widehat{\beta_{d,2}}(\xi_1, \xi_2) = \left( \text{sinc} \left( \frac{\xi_1}{2} \right) \text{sinc} \left( \frac{\xi_2}{2} \right) \right)^{d+1}$$

has its first zero at  $\xi_1, \xi_2 = \pm 2\pi$ , that is twice the Nyquist frequency corresponding to the integer sampling grid. The Fourier transform can even have significant contents outside of the square  $[-2\pi, 2\pi]^2$  depending on the spline degree  $d$ , this is why we must take into account values of the MTF (*Modulation Transfer Function*) outside the Nyquist square  $[-\pi, \pi]^2$ .

Another drawback of splines in the deblurring case is the periodic convolution of a spline function in  $S_{per}^d(n, m)$  with a kernel  $h \in L^2(\Omega)$  which does not necessarily lie in the same spline space or in any other spline space  $S_{per}^d(n, m)$ . Though, two cases indeed verify this property, the sensor PSF ( $h_{sens}$ ) without the exponential term and the movement equivalent PSF ( $h_{mov}$ ) described in section 4.1.1 :

$$\begin{cases} h_{sens,1}(x, y) = \mathbb{1}_{[-1,1]^2}(x, y) = \beta_{0,2} \left( \frac{x}{2}, \frac{y}{2} \right) \\ h_{mov}(x, y) = \left( \mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]} \otimes \delta_0 \right) (x, y) = (\beta_0 \otimes \delta_0) (x, y) \end{cases} \quad (4.21)$$

$$\Rightarrow \begin{cases} h_{sens,1} * \beta_{d,2}(\cdot) = ((\delta_0 + \delta_{-1}) \otimes (\delta_0 + \delta_{-1})) * \beta_{(d+1),2} \left( \cdot - \left( \frac{1}{2}, \frac{1}{2} \right) \right) \\ h_{mov} * \beta_{d,2}(\cdot) = (\beta_{d+1} \otimes \beta_d) (\cdot) \end{cases} \quad (4.22)$$

The optical PSF  $h_{opt}$  is generally known only in the Fourier domain ( $\hat{h}_{opt}$ ) and does not belong to a spline space. Given  $u \in S_{per}^d(n, m)$ , the convolution product  $h_{opt} * u$  rewrites as a combination of non-spline functions :

$$h_{opt} * u(\cdot) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \phi(\cdot - (k, l))$$

where  $\phi = h * \beta_{3,2}$  may have a very large support depending on  $h_{opt}$ . The sampling operator of these non-spline functions loses the nice numerical property of splines and no longer has a small number of nonzero coefficients. A solution to reduce the number of coefficients is the truncation of  $\phi$  but the exact computation of this function remains problematic since it is not piecewise polynomial.

The solution we propose here is an approximation of  $\phi$  in the  $L^2$  sense without truncation. In this projection process, the convolution operator on  $\beta_{d,2}$  is replaced by a convolution operator on the spline coefficients thanks to the periodic structure of  $S_{per}^d(n, m)$ . We consider here a very general function  $h \in L^2(\Omega)$  defined by its Fourier coefficients, but cases when  $\phi = h * \beta_{d,2}$  has small support and can be computed exactly should not be treated by this approach (this is the case for  $h_{sens,1}$  and  $h_{mov}$ ).

We define the projection problem (4.23) related to convolution operator  $h$  and spline spaces  $S_{per}^d(n, m)$  ( $d \in \{d_1, d_2\}$ ) where  $d_1$  is the degree of the approximation space for  $u$  (same as before) and  $d_2$  is the degree of the approximation space for  $\phi$  :

$$\text{proj}_{S_{per}^{d_2}(n, m)}(h * \beta_{d_1, 2}) = \min_{H \in \mathbb{R}^{n \times m}} \left\| h * \beta_{d_1, 2} - \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} H_{k,l} \beta_{d_2, 2}(\cdot - (k, l)) \right\|_{L^2(\Omega)}^2 \quad (4.23)$$

This problem admits a unique solution by the projection theorem and because the periodic functions  $\beta_{d_2, 2}$  form a basis of the spline space  $S_{per}^{d_2}(n, m)$ . It amounts to solve a linear system involving scalar products between splines and/or the convolution kernel  $h$  :

$$\sum_{k=0}^{n-1} \sum_{l=0}^{m-1} H(k, l) \langle \beta_{d_2, 2}(\cdot - (k, l)), \beta_{d_2, 2}(\cdot - (q, p)) \rangle = \langle h * \beta_{d_1, 2}, \beta_{d_2, 2}(\cdot - (q, p)) \rangle \quad \forall q, p \quad (4.24)$$

The scalar products between splines can be easily computed by formula (2.20) and the other scalar products can also be computed exactly by Parseval's formula :

$$c_{q,p} = \langle h * \beta_{d_1, 2}, \beta_{d_2, 2}(\cdot - (q, p)) \rangle_{L^2(\Omega)} = (mn) \langle \mathcal{F}(h * \beta_{d_1, 2}), \mathcal{F}(\beta_{d_2, 2}(\cdot - (q, p))) \rangle_{l^2(\mathbb{Z})} \quad (4.25)$$

where the discrete scalar product (right term) must be computed on the grid  $\frac{2\pi}{n}\mathbb{Z} \times \frac{2\pi}{m}\mathbb{Z}$ . The right term rewrites :

$$\mathcal{F}(h * \beta_{d_1, 2}) = \frac{1}{(mn)^2} \mathcal{F}(h) \mathcal{F}(\beta_{d_1, 2})$$

and we get a formulation of scalar products ( $c_{q,p}$ ) as a series "similar" to a Riemann series :

$$c_{q,p} = \frac{1}{mn} \sum_{k,l \in \mathbb{Z}^2} [\mathcal{F}(h) \mathcal{F}(\beta_{d_1, 2}) \mathcal{F}(\beta_{d_2, 2})] \left( \frac{2k\pi}{n}, \frac{2l\pi}{m} \right) e^{i(q(\frac{2k\pi}{n}) + p(\frac{2l\pi}{m}))} \quad (4.26)$$

Of course we cannot compute this infinite series but it converges very fast for usual splines degrees  $d_1, d_2 \approx 3$  since the FTM  $\hat{h}$  is bounded by 1 (and decreases exponentially fast for  $h_{opt}$ ) and the sinc term decreases as a  $O(\frac{1}{\|\xi_1 \xi_2\|^{d_1 + d_2 + 2}})$ . In practice the computation of this series for frequencies in  $[-\kappa\pi, \kappa\pi]^2$  with  $\kappa = 2$  is sufficient as shown in figure (4.9). Accurate computation of the series

in (4.25) is possible up to any precision in a reasonable time thanks to fast Fourier transform on  $[-\kappa\pi, \kappa\pi]^2$  and sub-sampling (or equivalently a periodization in the Fourier domain) :

$$c_{q,p} = \mathcal{F}^{-1} \left( [\mathcal{F}(h) \mathcal{F}(\beta_{d_1,2}) \mathcal{F}(\beta_{d_2,2})] \left( \frac{2k\pi}{n}, \frac{2l\pi}{m} \right) \right) (\kappa q, \kappa p)$$

with a  $K n \times \kappa m$  signal.

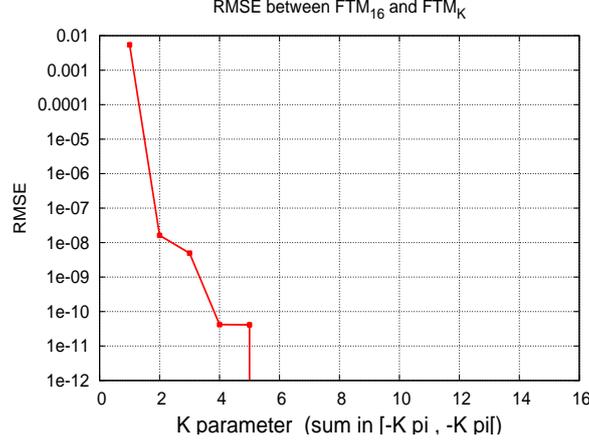


Figure 4.9: Mean Square Error between the result of (4.26) at given  $\kappa$  and the result at  $\kappa = 16$ . The error drops below the single precision for  $\kappa = 6$  and is already very low for  $\kappa = 2$ . (here we chose  $d_1 = d_2 = 3$ )

Finally the resolution of (4.24) is easily performed by means of FFT since the involved matrix is circulant with corresponding filter  $b$  :

$$\begin{aligned} M_{(k,l),(q,p)} &= \langle \beta_{d_2,2}(\cdot - (k,l)), \beta_{d_2,2}(\cdot - (q,p)) \rangle \\ &= \beta_{(2d_2+1),2}((k,l) - (q,p)) \\ &= b[(k-q, l-p)] \end{aligned} \quad (4.27)$$

$$\Rightarrow \mathcal{F}(H) = \frac{\mathcal{F}(c)}{\mathcal{F}(b)}$$

The division is well defined in the Fourier domain since  $|\mathcal{F}(b)(\xi)| > 0$  for all spline order  $d_2$  and all  $\xi \in [-\pi, \pi]$ .

This projection finally allows the computation of an approximation of  $h_{opt} * \beta_{d_1,2}$  with a spline sampling operator and a discrete convolution on splines coefficients :

$$\begin{aligned} \text{proj}_{S=S_{p\tilde{e}_r}^{d_2}(n,m)}(h * u) &= \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \text{proj}_S(h * \beta_{d_2,2}(\cdot - (k,l))) \\ &= \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \sum_{q=0}^{n-1} \sum_{p=0}^{m-1} H_{q,p} \beta_{d_2,2}(\cdot - (k-q, l-p)) \\ &= \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} a_{k,l} \sum_{q'=0}^{n-1} \sum_{p'=0}^{m-1} H_{k-q', l-p'} \beta_{d_2,2}(\cdot - (q', p')) \quad (\text{change of indices}) \\ &= \sum_{q'=0}^{n-1} \sum_{p'=0}^{m-1} \left( a * \check{H} \right) [q', p'] \beta_{d_2,2}(\cdot - (q', p')) \quad (\text{invert summations}) \end{aligned} \quad (4.28)$$

We will use this projection in the deblurring problem with a non-spline convolution kernel and show that it preserves the good approximation property of splines.

#### 4.2.3.2 Consistency with Shannon's conditions

We know how to model blurry data by means of convolution kernels, but what is exactly restoration of blurry data ? In the simple regular sampling case, a blurry image is modelled as :

$$v[k, l] = (u * h)[k, l] + w[k, l] \quad (4.29)$$

The inverse blurring operator  $h^{-1}$  such that  $\mathcal{F}(h^{-1}) = \frac{1}{\mathcal{F}(h)}$  can be numerically unstable and the application of  $h^{-1}$  produces coloured noise  $h^{-1} * w$  with possible high Fourier coefficients. This is why a trade-off between regularity and data fitting is necessary (by means of local or non-local regularizations) and leads to formulations of the type :

$$u^* = \underset{u}{\operatorname{argmin}} \frac{1}{2} \|h * u - v\|^2 + E(u) \quad (4.30)$$

This data-fitting term means that we want to recover a fully deblurred image, while forgetting the most important theorem in image sampling : Shannon's theorem. For a given sampling step  $\Delta$  the sampled image  $(u(k\Delta, l\Delta))_{(k,l) \in \mathbb{Z}^2}$  has Fourier transform proportional to the periodization of  $\hat{u}$  with periods  $(\frac{2\pi}{\Delta}, 0)$  and  $(0, \frac{2\pi}{\Delta})$ .

$$\mathcal{F}(u(k\Delta, l\Delta)) = \sum_{q \in \mathbb{Z}} \sum_{p \in \mathbb{Z}} \hat{u}(\xi_1 + q\frac{2\pi}{\Delta}, \xi_2 + p\frac{2\pi}{\Delta})$$

An image generally has important contents at low frequencies and if  $\hat{u}$  has large contents outside the Nyquist square  $[-\frac{\pi}{\Delta}, \frac{\pi}{\Delta}]$ , the periodization in frequency domain will alter the frequency contents. This well known phenomenon called *aliasing* has a great influence in sampling design. In order to acquire non-aliased images, the device MTF has low amplitudes at the boundaries of the Nyquist square  $[-\frac{\pi}{\Delta}, \frac{\pi}{\Delta}]$ . Thus, the inversion of  $h$  in (4.29) cancels the frequency attenuation of the MTF which allowed non-aliased acquisition.

The user is the only judge for the accuracy of the restoration since there does not exist a single reference image for this problem. Indeed, every image sampled with frequency  $\frac{1}{\Delta}$  and previously convolved with a kernel  $h_{ref}$  respecting Shannon's conditions is a valid image. This curse cannot be avoided unless we address the super-resolution or image fusion problems.

In all our experiments, the reference image is obtained by convolution of a high-resolution image  $u_0$  with a 2D *prolate kernel*  $p_\alpha$  and sub-sampling according to the prolate cut-off frequency. A prolate function as defined by Landau, Pollak and Slepian [85] is a time-limited signal (small support filter here) and has most of its energy concentrated in  $[-\alpha\pi, \alpha\pi]$  in the Fourier domain. The factor  $\alpha \in ]0, 1]$  is the cut-off parameter although the Fourier coefficients of the prolate are not necessarily zero above  $\alpha\pi$ . The construction of 2D prolates is similar to the 1D case and is performed by iterative Fourier projections with the operator :

$$\mathcal{F}^{-1} \circ P_\alpha \circ \mathcal{F}$$

where  $P_\alpha$  is the support restriction in Fourier domain (it can be a square with size  $2\pi\alpha$  or disc with radius  $\alpha\pi$ ). The cut-off parameter is chosen as the inverse of the sub-sampling parameter, and the prolate has generally a support length about 3 – 5. The images produced this way are visually well sampled, i.e. high frequency structures do not seem aliased by this filtering+sub-sampling procedure. The deblurring should be aimed at recovering the image  $u_0 * p_\alpha$  on the coarse grid

$\Lambda_{reg}$  of the sub-sampling step, and there exist at least two ways to achieve this goal. The first one consists in a new formulation of the deblurring problem :

$$u^* = \underset{u}{\operatorname{argmin}} \quad \frac{1}{2} \|(h * p_\alpha^{-1}) * u - v\|^2 + E(u) \quad (4.31)$$

but the numerical behaviour of the convolution operator  $h * p_\alpha^{-1}$  is not well defined since the prolate function can take arbitrary low values in the Fourier domain. Furthermore, this implies that the solution  $u^*$  of (4.31) depends on  $p_\alpha$ , the second approach we propose is more flexible. Consider  $v^*$  the solution of the initial deblurring problem and  $u^*$  defined by :

$$v^* = \underset{u}{\operatorname{argmin}} \quad \frac{1}{2} \|(h * u) - v\|^2 + E(u) \quad , \quad u^* = p_\alpha * v^* \quad (4.32)$$

Here the regularization should only prevent the noise explosion. The aliasing problem may appear in  $v^*$  (depending on the regularization) but it is no more problematic because  $u^*$  is obtained by a further convolution of  $v^*$  with a low-pass filter. This way we can consider the usual deblurring problem (with  $h$ ) and compare our results to the reference we want. The prolate kernel may even be replaced by a more general low-pass filter, with a well-chosen value at the Nyquist frequency.

Finally we present the deblurring formulation we use in the following for the irregular sampling problem :

$$\boxed{v^* = \underset{u \in S_{per}^{d_1}(n,m)}{\operatorname{argmin}} \quad \frac{1}{2} \left\| \underset{S_{per}^{d_2}(n,m)}{\operatorname{proj}} (h * u)(x_k, y_k) - v_k \right\|^2 + \lambda TV_{\mathbb{Z}}(u) \quad , \quad u^* = p_\alpha * v^*} \quad (4.33)$$

and the reference we choose (see figure 4.10) :

$$\boxed{u_{ref} = \left( u_0 * p_{\left(\frac{1}{K}\right)} \right) \downarrow_{\kappa}}$$

$K$  is typically 2 or 3.

We propose to use the classical TV on the integer grid, as a deblurring and denoising regularization with a global constraint. We know from [4] that TV produces stair-casing effect for high regularizations, but here the noise level is usually low ( $\sigma \approx 1$  but we tested higher noise levels) and the post-filtering step should be taken into account as it will smooth the sharp edges of TV and decrease the remaining noise.

#### 4.2.3.3 Variational formulation

We rewrite the function in (4.33) as a function of the spline coefficients  $(a_{k,l})$  :

$$E(a) = \|S_\Lambda H_{proj} a - v\|^2 + \lambda E_{TV}(M_{d_1}(a)) \quad (4.34)$$

where  $H_{proj}$  is the convolution matrix with the 2D filter associated to the spline projection operator (4.28) and  $M_{d_1}$  is the linear operator associated to the inverse spline transform

$$(M_{d_1} a)_{k,l} = u(k, l)$$

For the moment we call  $E_{TV}(u)$  a general regularization function which behaves like TV, it is not necessarily differentiable so we consider the subdifferential inclusion related to equation 4.34.

$$0 \in (S_\Lambda H_{proj})^t (S_\Lambda H_{proj} a - v) + \lambda \partial(E_{TV} \circ M_{d_1})(a)$$

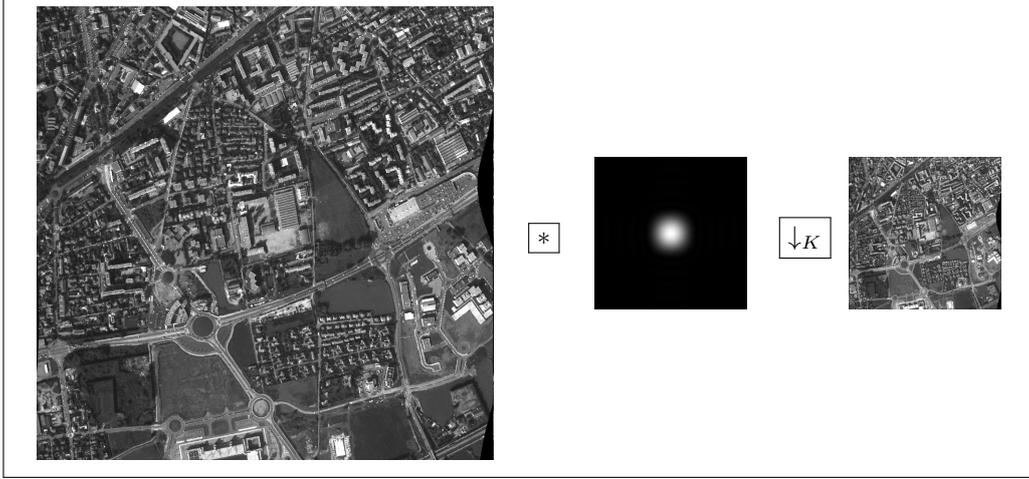


Figure 4.10: How we generate reference images for the deblurring problem : first a high resolution image (CNES copyrighted image) (left) is convolved with a low-pass prolate with cut-off frequency  $\frac{1}{K}$  (Fourier transform is displayed in the middle), it is then subsampled with a factor  $K \in \{2, 3, \dots\}$  (right). The result is visually nice and does not suffer from the aliasing effect.

In this form the problem looks like the ACT+TV formulation (see section 4.1.2) in which the authors compute the proximal of  $TV \circ Q^{-1/2}$ . Here the operator  $(M_{d_1})$  is always invertible, its inverse has a fast implementation for periodic signals by means of FFT, it also has a much faster implementation with FIR filters as described in [89] for other boundary conditions.

We adopt a different approach and reformulate the problem in terms of  $U = M_{d_1} a$  (the regularly sampled image) instead of  $(a_{k,l})$  the spline coefficients, this allows the direct computation of matrix-vector products with matrix  $H_{proj} M_{d_1}$  thanks to the Fourier transform since  $M_{d_1}$  is circulant in periodic spline spaces.

$$\begin{aligned} V^* &= \operatorname{argmin}_{U \in \mathbb{R}^{m \times n}} E(U) \\ &= \operatorname{argmin}_{U \in \mathbb{R}^{m \times n}} \|S_\Lambda H_{proj} (M_{d_1})^{-1} U - v\|^2 + \lambda E_{TV}(U) \end{aligned} \quad (4.35)$$

The two minimization problems (4.34) and (4.35) are equivalent but the subdifferential inclusion associated to (4.35) now uses a simpler subdifferential operator and we can compute matrix-vector products  $H_{proj} (M_{d_1})^{-1} U$  in the Fourier domain with fast FFT with a cost of 2 FFTs and one multiplication on complex data.

$$0 \in (M_{d_1})^{-t} H_{proj}^t S_\Lambda^t (S_\Lambda H_{proj} (M_{d_1})^{-1} U - v) + \lambda \partial E_{TV}(U)$$

$$H_{proj} (M_{d_1})^{-1} U \asymp O((m \times n) \log m \times n) + 2(m \times n)$$

We can use the Forward-Backward algorithm and accelerated variants to solve problem (4.35) as described in algorithm 4.11.

As in previous chapter, we need to compute iteratively a proximal operator of a TV-like energy. Next section provides details of this computation for nonsmooth and smooth TV with a dual approach.

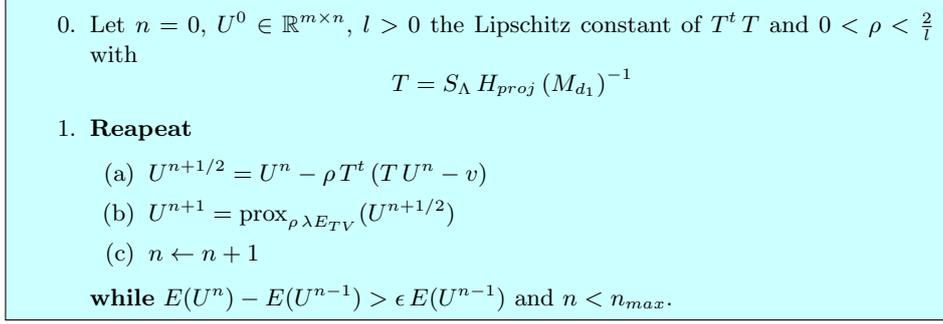


Figure 4.11: Forward-Backward algorithm for the resampling+deblurring problem (4.35)

#### 4.2.3.4 Proximal of TV

The TV-denoising problem (i.e. proximal of TV) has a very dense literature, indeed many authors have contributed to this non-quadratic minimization problem in both discrete and continuous case, among whom Chambolle et al. [22] [21] [24] (2010).

Several TV-like energies have emerged, here we consider the classical TV :

$$TV(U) = \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} \|\nabla_{++} U(k, l)\|$$

where  $\nabla_{++} : \mathbb{R}^{m \times n} \rightarrow (\mathbb{R}^2)^{m \times n}$  is the finite differences operator defined by :

$$\begin{aligned} (\nabla_{++} U)_1(k, l) &= \begin{cases} U(k+1, l) - U(k, l) & \text{if } k \neq n-1 \\ 0 & \text{else} \end{cases} \\ (\nabla_{++} U)_2(k, l) &= \begin{cases} U(k, l+1) - U(k, l) & \text{if } l \neq m-1 \\ 0 & \text{else} \end{cases} \end{aligned} \quad (4.36)$$

This operator is the classical one in image processing, we will use the simpler  $\nabla$  notation from now. The classical TV is convex (and thus continuous), positive homogeneous and its conjugate is the indicator function of a convex set :

$$TV^* = \mathbb{I}_{\nabla^* K_{TV}} \quad , \quad K_{TV} = \{s \in (\mathbb{R}^2)^{m \times n} \text{ s.t. } \|s_{i,j}\| \leq 1 \text{ for all } i, j\} \quad (4.37)$$

Thus the  $\rho$ -proximal of this TV energy amount to compute the projection of a vector  $U$  on the convex set  $\rho \nabla^* K_{TV}$  and the result is given by :

$$\text{prox}_{\rho TV}(U) = U - \text{proj}_{\rho \nabla^* K_{TV}}(U) \quad (4.38)$$

A smooth TV energy with the same nice property as studied in [72] derives from the Huber function  $\psi_\mu : \mathbb{R} \rightarrow \mathbb{R}$  :

$$TV_\mu(U) = \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} \psi_\mu(\|\nabla U(k, l)\|) \quad , \quad \psi_\mu(x) = \begin{cases} |x| & \text{if } |x| \geq \mu \\ \frac{x^2}{2\mu} + \frac{\mu}{2} & \text{else} \end{cases} \quad (4.39)$$

Huber's function rewrites as a supremum :

$$\psi_\mu(x) = \frac{\mu}{2} + \sup_{|s| \leq 1} x s - \frac{\mu}{2} s^2$$

thus, its conjugate is  $\psi_\mu^*(s) = \frac{\mu}{2}(s^2 - 1) + \mathbb{I}_{[-1,1]}(s)$ .

As  $TV_\mu = \sum_{k,l} \psi_\mu(\|\nabla U(k,l)\|)$ , we can rewrite  $TV_\mu$  as a supremum :

$$\begin{aligned} TV_\mu(U) &= \sup_{S \in [-1,1]^{m \times n}} \sum_{k,l} \|\nabla U(k,l)\| S_{k,l} - \frac{\mu}{2} \sum_{k,l} (S_{k,l}^2 - 1) \\ &= \sup_{S' \in K_{TV}} \langle \nabla U, S' \rangle - \frac{\mu}{2} (\|S'\|^2 - mn) \end{aligned}$$

where the scalar product on  $\langle \nabla U, S' \rangle$  is equal to  $\sum_{k,l} ((\nabla U)_1(k,l)(S')_1(k,l) + (\nabla U)_2(k,l)(S')_2(k,l))$ . The inf-sup exchange provides the expression of the proximal of  $TV_\mu$  :

$$\begin{aligned} &\sup_{S' \in K_{TV}} \min_{V \in \mathbb{R}^{m \times n}} \lambda \left( \langle \nabla V, S' \rangle - \frac{\mu}{2} (\|S'\|^2 - mn) \right) + \frac{1}{2} \|U - V\|^2 \quad (\mathcal{D}_{TV_\mu}) \\ \Leftrightarrow &V = U - \lambda \nabla^* S' \quad \text{and} \quad S' = \operatorname{argmin}_{T \in K_{TV}} \frac{1}{2} \|\nabla^* T - \frac{1}{\lambda} V\|^2 + \frac{\mu}{2\lambda} \|T\|^2 \quad (4.40) \end{aligned}$$

The proximal of  $TV_\mu$  is thus as easy to compute as the proximal of classical TV, which corresponds to the particular case  $\mu = 0$ . Problem (4.40) can be solved by the same algorithms as classical TV dual approaches (see [24] for a review). For the classical TV proximal we use Chambolles's algorithm in [22] although more efficient schemes now exist : FB, mFISTA, Nesterov, all these algorithms can be applied to this problem.

- Let  $S^0 \in K_{TV}$ ,  $l > 0$  the Lipschitz constant of  $\nabla \nabla^*$  and  $0 < \tau \leq 0.25$ .
- $g^n = \nabla (\nabla^* T^n - \frac{1}{\lambda} V)$
- $(S^{n+1})_{k,l} = \frac{1}{1 + \tau \|(g^n)_{k,l}\|} (S^n + \tau g^n)_{k,l}$  for all  $k, l$ .

Figure 4.12: Chambolle's algorithm for the dual minimization problem (4.40) with  $\mu = 0$  (classical TV).

On the other hand we solved problem (4.40) by the projected gradient algorithm (which coincides with the Forward-Backward and can also be applied to classical TV) :

- Let  $S^0 \in K_{TV}$ ,  $l > 0$  the Lipschitz constant of  $(\nabla \nabla^* + \frac{\mu}{\lambda} Id)$  and  $0 < \tau < \frac{2}{l}$ .
- $g^n = \nabla (\nabla^* S^n - \frac{1}{\lambda} V) + \frac{\mu}{\lambda} S^n$
- $S^{n+1} = \operatorname{proj}_{K_{TV}} (S^n - \tau g^n)$

Figure 4.13: Projected gradient for the minimization of the dual problem associated to the proximal of  $TV_\mu$ .

More recently the authors of [24] have compared several algorithms for the TV- $l^2$  and show some experiments in which accelerated primal-dual algorithms converge faster than the FB, mFISTA and Nesterov's algorithms, works of this section could certainly benefit from these breakthroughs. Figure (4.14) below shows the tuning of the step parameter  $\tau$  in the projected gradient algorithm

4.13. Chambolle's algorithm is known for its monotonicity, the primal energy is strictly decreasing for  $0 < \tau \leq \frac{1}{4}$ , unfortunately this is not the case for the projected gradient if  $\tau \approx \frac{2}{l}$  with  $l$  the Lipschitz constant of  $\nabla \nabla^* + \frac{\mu}{\gamma} Id$ .

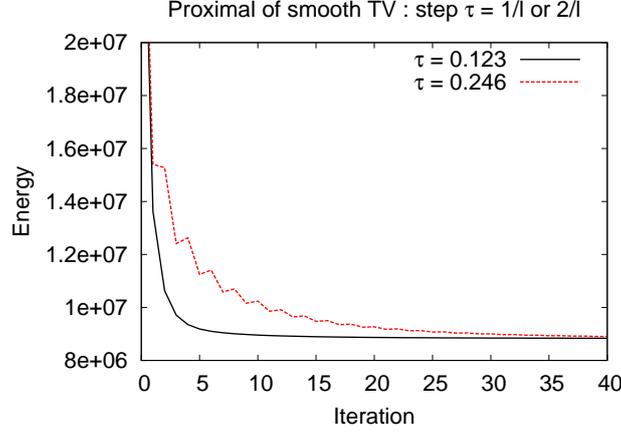


Figure 4.14: Energy in the proximal computation of  $TV_\mu$  by the projected gradient algorithm. The energy is not guaranteed to decrease for steps  $\tau$  close to the theoretical bound  $\frac{2}{l}$ , one should care about this when using the smooth TV proximal.

It is thus possible to use a smooth TV energy as a regularization for our problem, but the post-filtering of  $V^*$  in our approach already reduces the stair-casing effect, this is why we mainly used the classical TV function.

#### 4.2.3.5 The whole algorithm

The function we minimize in (4.35) is a classical Lagrangian for the constraint :

$$\|S_\Lambda H_{proj} M_{d_1}^{-1} U - v\|^2 \leq \theta |\Lambda| \sigma^2$$

with  $\theta$  is the same partial denoising parameter as in (4.17) p. 86. We consider the same update of Lagrange parameter  $\lambda$  in the Uzawa's loop, that is :

$$\lambda_{n+1} = \lambda_n / c(V_n^*) \quad (4.41)$$

where  $V_n^*$  is the result obtained by the Forward-Backward algorithm at step  $n$  with  $\lambda_n$  and  $c(\cdot)$  is defined by :

$$c(V) = \exp\left(\frac{\|S_\Lambda H_{proj} M_{d_1}^{-1} V - I\|^2 - R |\Lambda| \sigma^2}{R |\Lambda| \sigma^2}\right)$$

with minimum value  $c_{min}$  and maximum value  $c_{max}$  two fixed parameters controlling the Lagrangian update in cases of bad initialization. This is summed up in algorithm 4.16.



Figure 4.15: Experiments with the smooth TV proximal (denoising). For low smoothing parameters the difference is hardly visible ((b) vs (c)), but for higher smoothness the staircasing effect is clearly reduced ((b) vs (d)).

0. Let  $n = 0$ ,  $R \in ]0, 1]$ ,  $W \in ]0, R[$ .

1. Compute initial guess  $V_{-1}^* \in \mathbb{R}^{n \times m}$ ,  $\lambda_0 > 0$  and  $l > 0$  the Lipschitz constant of  $T^* T$  with

$$T = S_\Lambda H_{proj} M_{d_1}^{-1}$$

2. **Repeat**

(a) Compute  $V_n^*$  by the Forward-Backward algorithm 4.11 with Lagrange parameter  $\lambda_n$ , stopping parameter  $\epsilon > 0$  and initial vector  $U^0 = V_{k-1}^*$ .

(b) Update  $\lambda_{n+1}$  by formula (4.41).

(c)  $n \leftarrow n + 1$

**While**  $\|T V_n^* - v\| \notin [(R - W) |\Lambda| \sigma^2, (R + W) |\Lambda| \sigma^2]$

Figure 4.16: Whole deblurring algorithm : we kept the stopping criterion and multiplicative update of the Lagrange parameter. The FB algorithm amounts to compute iteratively a proximal, this is done either by Chambolle's algorithm for classical TV energy or Projected Gradient (FB) for the smooth  $TV_\mu$ .

### 4.2.3.6 Experiments

We present below the set of images we used as a validation for our algorithm. These images have been obtained by irregular sampling of larger resolution images together with a low-pass filtering (the prolate for example) and a band-limited hypothesis. The irregularity in the sampling set is obtained by adding a low-frequency Gaussian variable to the regular integer coordinates as shown in figure 4.17.

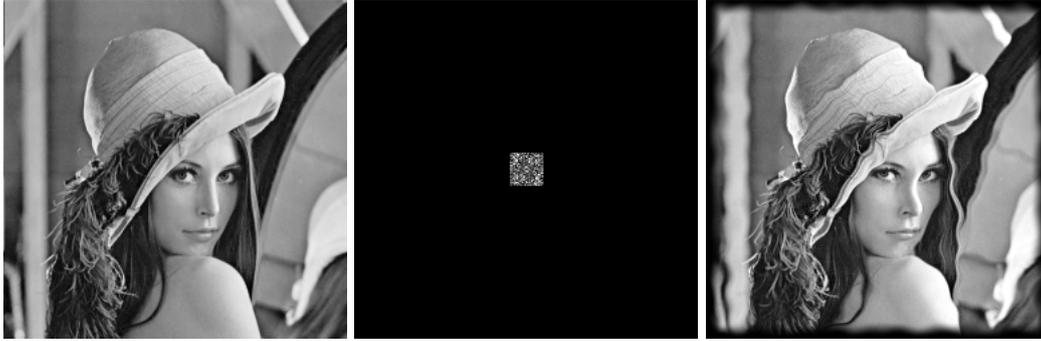


Figure 4.17: Original Lena image (left), perturbation in the Fourier domain (middle) and perturbed image (right). The border smoothing is aimed at avoiding periodization problem in the simulation and restoration and depends here on the maximum perturbation.

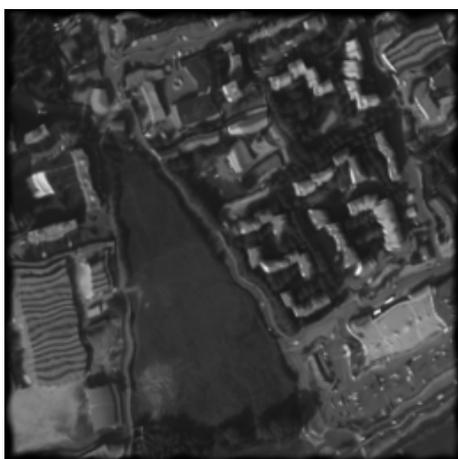
This is the kind of perturbations we can find in satellite images as modeled in [2]. The results we present were obtained with the comparison criterion described in section 4.2.3.3 and bicubic splines (the approximation space is thus  $S_{per}^3(n, m)$ ).

We used the non-spline kernel  $h_{opt} * h_{sens,2}$ , it is a strongly smoothing kernel with a singularity at the origin in the Fourier domain. It is due to the optical device blurring kernel ( $h_{opt}$ ) and to the diffusion of electrical charges between neighbour elements of the TDI ( $h_{sens,2}$ ). The results below correspond to the classical nonsmooth TV. The algorithm uses parameters  $\epsilon$  and  $\nu$  in the Forward-Backward and TV- $\ell^2$  problems, we used the value  $1.10^{-5}$  as a trade-off between accuracy and time.

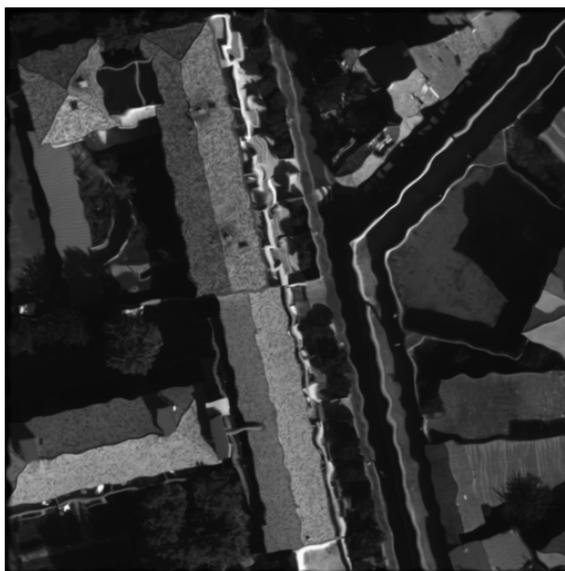
H]



(a) Beauvais image  $367 \times 590$



(b) Nîmes image  $257 \times 257$



(c) Stmichel image (Toulouse)  $512 \times 512$

Figure 4.18: Our test images have various resolutions : from coarsest (top) to finest (CNES copyrighted images) (bottom). Perturbations follow a low-frequency Gaussian law (with minimum period a tenth of the image sizes).

(a) PSNR measurements

image	noise level	ACT	FAR	CG-S	TIKH-S	TV-S
beauvais	1	24.3587	39.2606	38.6486	39.822	<b>39.9626</b>
	3	24.3408	33.2418	32.7089	33.5623	<b>33.9708</b>
	5	24.3132	30.9623	30.3513	31.1895	<b>31.4893</b>
	7	24.27	29.6102	29.051	29.5223	<b>29.8931</b>
nimes	1	26.6265	40.3505	38.816	39.8873	<b>40.9131</b>
	3	26.5929	34.2299	32.7798	33.7992	<b>34.9151</b>
	5	26.5374	31.9392	30.7082	31.4423	<b>32.3899</b>
	7	26.4442	<b>30.6132</b>	29.5112	29.8485	30.5806
stmichel	1	28.0076	42.7022	40.5794	41.5334	<b>43.1228</b>
	3	27.9677	36.9147	34.7294	35.8409	<b>37.5323</b>
	5	27.8996	<b>34.6547</b>	32.8676	33.3145	34.5249
	7	27.7711	<b>33.0287</b>	31.3372	31.119	32.4016

(b) Computation time

image	noise level	ACT	FAR	CG-S	TIKH-S	TV-S
beauvais	1	2.404s	24m36.808s	7.048s	28.430s	58.124s
	3	2.452s	8m55.637s	3.424s	12.461s	37.194s
	5	2.476s	5m55.774s	2.804s	11.749s	31.778s
	7	2.420s	4m36.645s	2.564s	10.893s	29.674s
nimes	1	0.828s	4m30.369s	1.180s	4.556s	12.169s
	3	0.836s	1m56.851s	0.652s	2.744s	9.201s
	5	0.816s	1m20.633s	0.552s	2.520s	8.273s
	7	0.824s	0m56.364s	0.516s	2.556s	9.453s
stmichel	1	2.860s	17m56.179s	2.888s	4.276s	34.522s
	3	2.852s	7m24.924s	1.732s	2.856s	28.926s
	5	2.920s	5m5.375s	1.588s	1.792s	33.410s
	7	2.872s	4m12.044s	1.404s	2.404s	30.506s

Table 4.1: PSNR (top) and computation time (bottom) for the deblurring experiments of images in data set 4.18. In terms of PSNR, our method (TV-S) usually provides very good results close to FAR. The adaptive frequency regularization in FAR depends on the image, we did not tune it here so it could be improved. PSNR should not be the only comparative criterion, we display extracts from the result of these methods in the figures below.

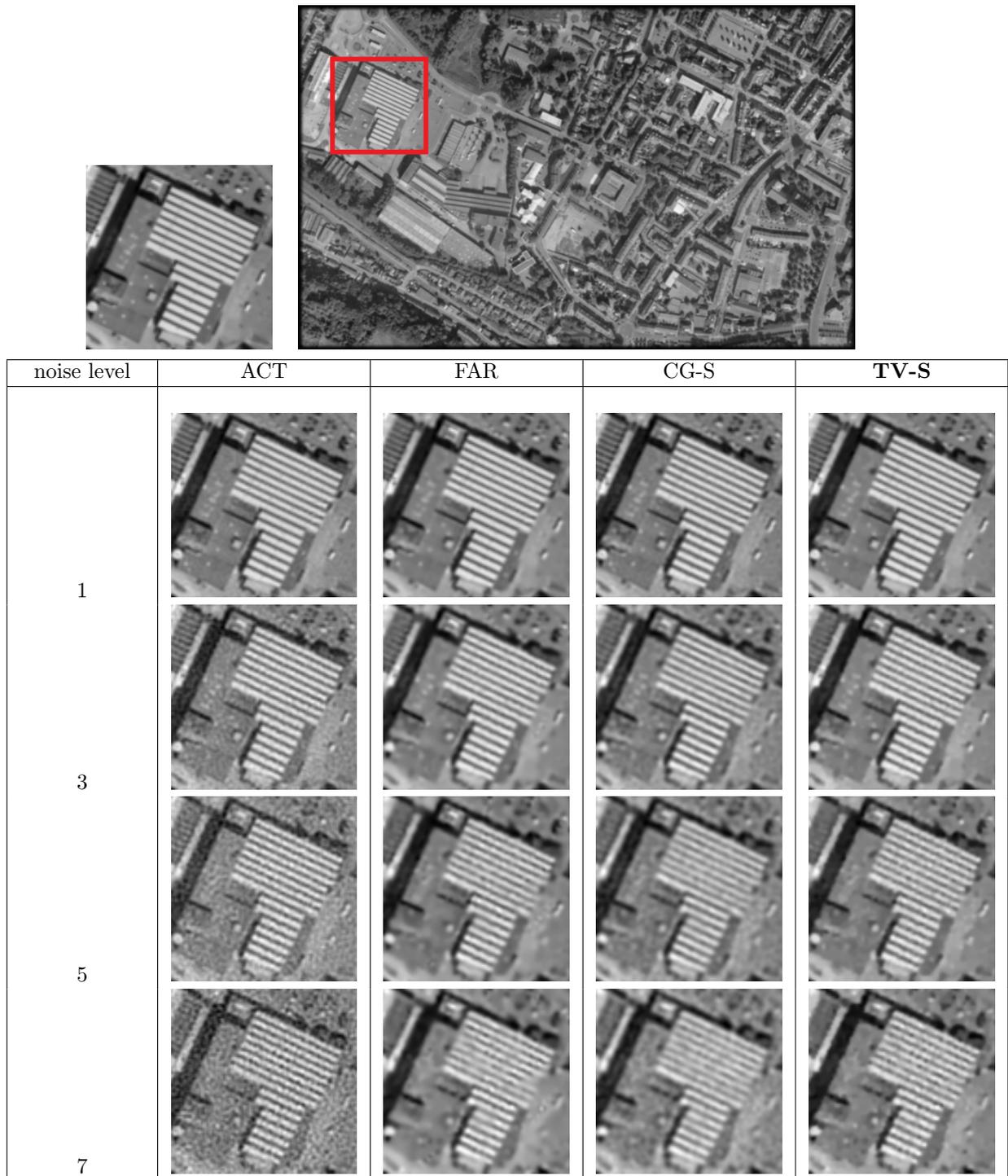


Figure 4.19: Extracts from the red square in top image. Despite the small stair-casing effect, TV regularization is sharper, it better denoises than other truncated methods (ACT and CG-S). Some textured areas in FAR and CG-S are not perfectly restored (it also appears in our method in a lesser extent), this might be due to a low regularization together with a low sampling density.

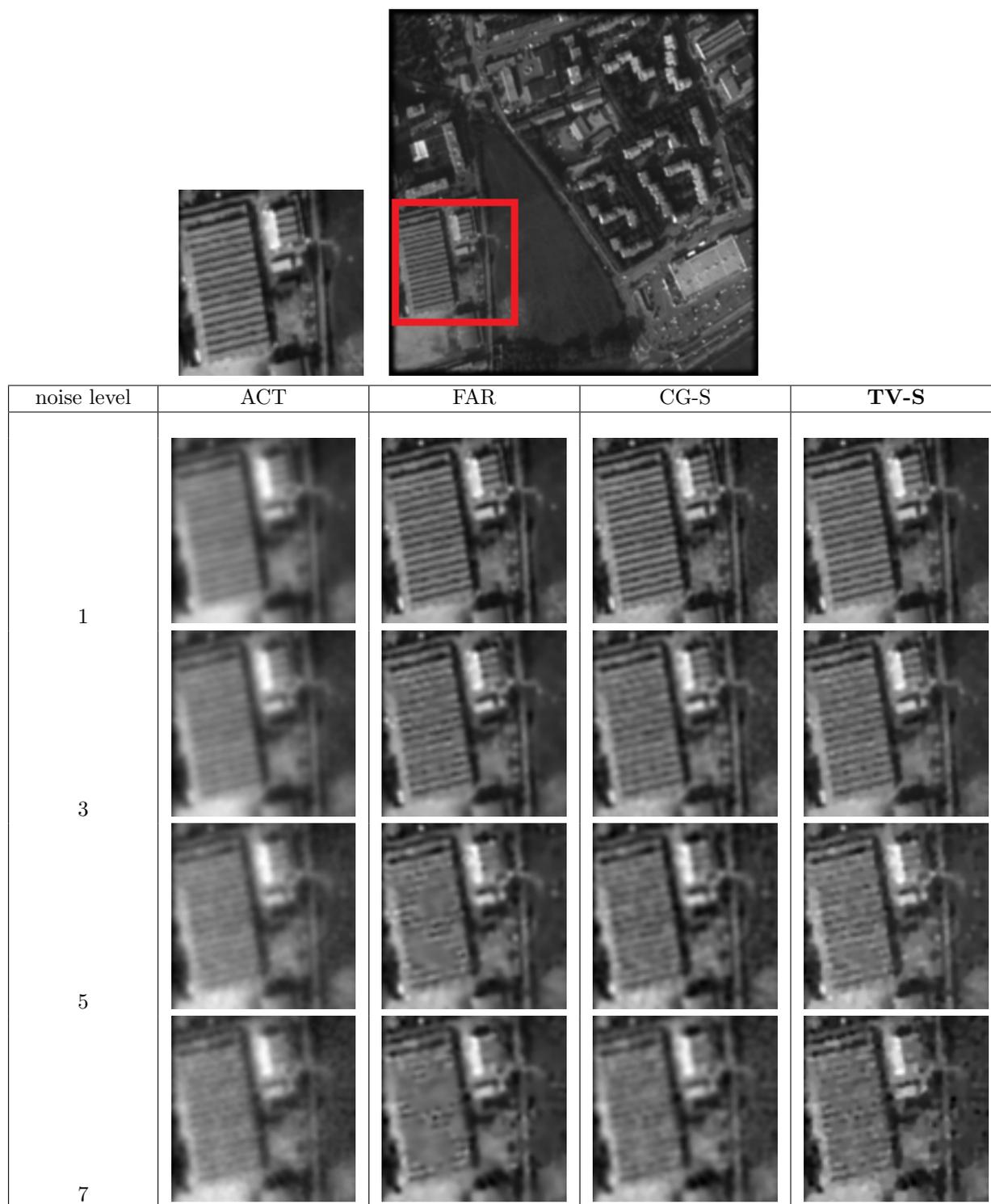


Figure 4.20: Extracts from image Nimes.

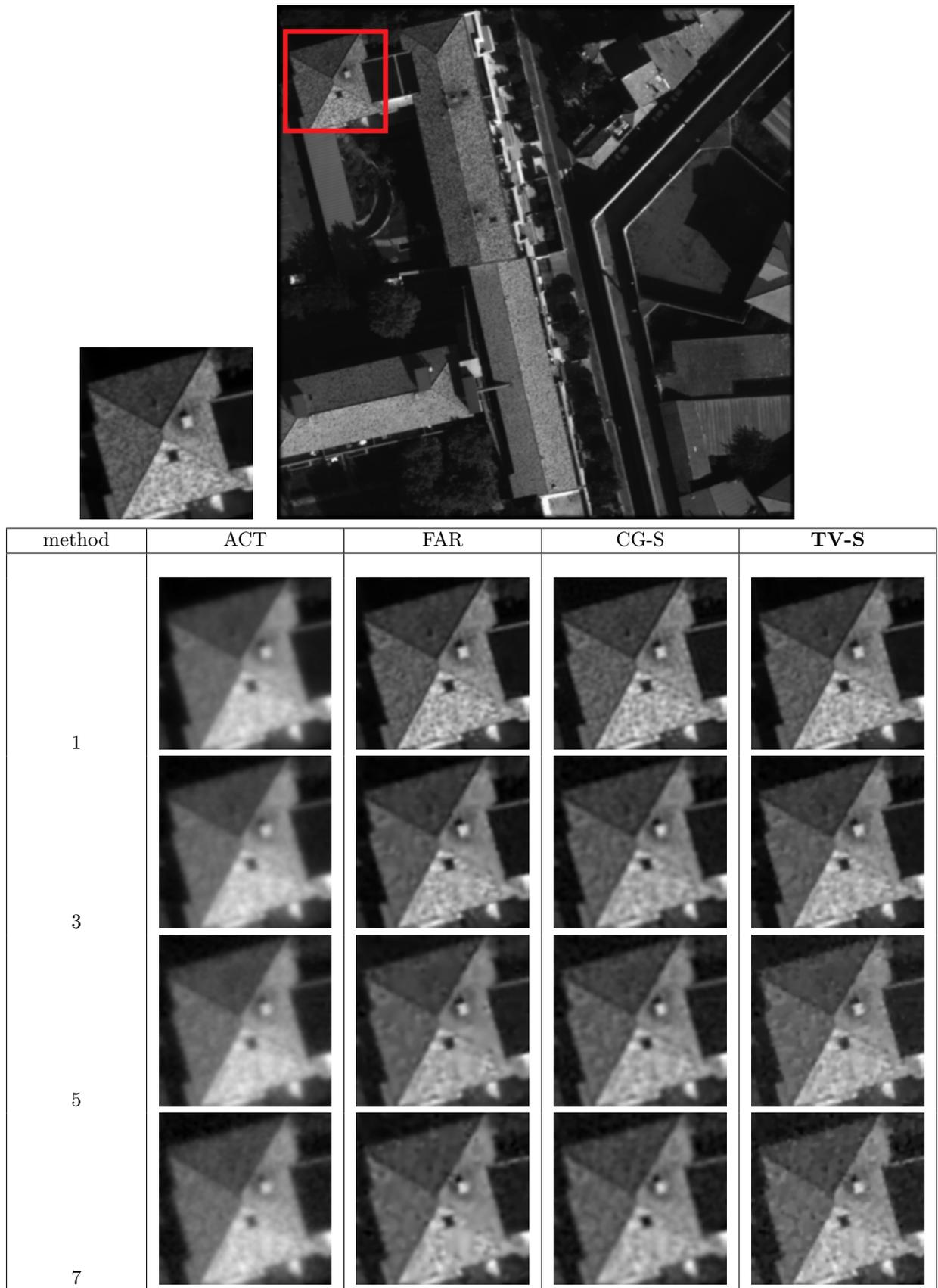


Figure 4.21: Extracts from image Stmichel.

image	noise level	ACT	FAR	CG-S	TIKH-S	TV-S
beauvais	1	32.6674 2.420s	45.2063 4m47.618s	44.7302 1.556s	<b>45.508</b> 4.800s	43.4553 14.809s
	3	32.5057 2.400s	38.647 2m4.076s	38.5967 1.284s	<b>39.6871</b> 2.408s	37.7864 13.005s
	5	32.2144 2.436s	35.4431 2m46.018s	35.3518 1.240s	<b>36.4137</b> 2.056s	34.8764 12.953s
	7	31.8291 2.412s	33.6107 3m24.809s	33.8758 1.268s	<b>34.2623</b> 2.148s	32.9373 12.721s
nimes	1	34.6487 0.836s	<b>46.4246</b> 0m39.454s	45.5707 0.380s	46.2529 0.976s	44.5179 3.372s
	3	34.3649 0.824s	39.4128 27.842s	38.8316 0.332s	<b>39.589</b> 0.628s	38.6841 3.216s
	5	33.7547 0.828s	36.2019 36.962s	35.0681 0.300s	<b>36.2501</b> 0.500s	35.8202 3.404s
	7	33.0592 0.828s	<b>34.3023</b> 46.167s	33.5249 0.320s	33.794 0.548s	33.9268 3.172s
stmichel	1	37.3659 2.896s	<b>48.0765</b> 2m33.406s	47.1979 1.248s	47.6472 4.276s	46.6596 12.209s
	3	36.8503 2.856s	<b>41.3718</b> 2m37.018s	39.6832 1.024s	40.8319 2.856s	40.5863 15.405s
	5	35.9848 2.932s	<b>38.3882</b> 3m32.937s	35.2134 0.932s	37.5416 1.792s	37.9242 16.609s
	7	34.9904 2.896s	<b>36.6007</b> 3m47.030s	33.9408 0.936s	35.0942 2.404s	36.1572 19.589s

Table 4.2: Results from the denoising problem (no deblurring). Quadratic regularization on the Hessian (TIKH-S) gives very good results with a very low computation time. FAR method is very close and our TV regularization with splines is worse than some unregularized methods (truncated CG with trigonometric polynomials (ACT) and the same with splines (CG-S)). TV regularization on the integer grid is not well-adapted to the resampling problem, the stair-casing effect is too strong.

#### 4.2.3.7 Observations

The goal of this formulation with splines was the reduction of the computation time and the preservation of the image quality. The first point is verified, compared to FAR (Quasi-Newton algorithm with smooth TV-like regularization at half-integers and local constraints) our algorithm is approximately 20 times faster for low noise levels and about 10 times faster for very high noise levels. We observed that the local constraints approach requires more iterations in Uzawa's loop (the update of Lagrange parameters) which explains a part of the computation time reduction, but this accounts only for a factor two at most for low noise levels, and the overall gain in computation time for typical image sizes seems to be a factor 10.

We obtained equivalent PSNR results but the reason for better looking results with global constraints (except for the stair-casing) compared to local constraints is unclear. As we can see in extracts from figures 4.19 and 4.20, the two methods have different behaviours in the presence of highly oscillating textures. In the FAR results, the textures are less regularized because of the local adaptation but the result looks smooth where the sampling density is low. Global regularization certainly has drawbacks (oversmoothing effect in textured regions) but it seems to have better inpainting properties in regions of low sampling density.

The stair-casing effect is present in our results, despite the post-processing step. This point could be improved, for example by using a smooth version of TV ( $TV_\mu$ ) or a different discretization of TV on the half-integers grid as described in the perspectives.

#### 4.2.3.8 Influence of some parameters

We provide here a short study of the influence of spline approximation degrees ( $d_1$  for the image and  $d_2$  for the projection) and smoothing parameter in  $TV_\mu$ .

The following experiments consist in the resolution of both problems ( $\mathcal{P}_{smp}$ ) and ( $\mathcal{P}_{blur}$ ) with arbitrary spline orders, although we kept the same approximation and projection degree convention  $d_1 = d_2$ . The overall Uzawa's algorithm structure is the same, as a consequence the regularization parameters may slightly vary with the approximation degree  $d_1$ . Table 4.3 shows the evolution of computation time and PSNR in the two algorithms. We also display the results in figure 4.22.

spline degree	PSNR and Computation time	
	resampling	deblurring
0	21.4673 45.623s	29.2966 1m57.731s
1	37.9543 6.852s	38.7212 44.299s
2	43.5424 4.336s	40.8748 18.453s
3	43.9713 6.252s	40.9 19.853s
4	44.0268 8.389s	40.851 21.937s
5	44.1683 9.745s	40.8517 16.969s
6	<b>44.2732</b> 14.853s	41.0533 22.933s
7	44.2308 22.053s	<b>41.1621</b> 25.390s

Table 4.3: Experiments on Nimes ( $259 \times 259$ ) image with different spline degrees. The zero order splines (piecewise constant functions) give very poor results, the Uzawa's loop fails in this case to find the appropriate Lagrange parameter (infeasible constraint). In all the other cases the algorithm performs well and the best results are obtained for high spline degrees (here 6 and 7). This is consistent with the image deformation model (trigonometric polynomials) and we observe that computation time is not very sensitive to the spline degree since most of the computation concerns the TV proximal.

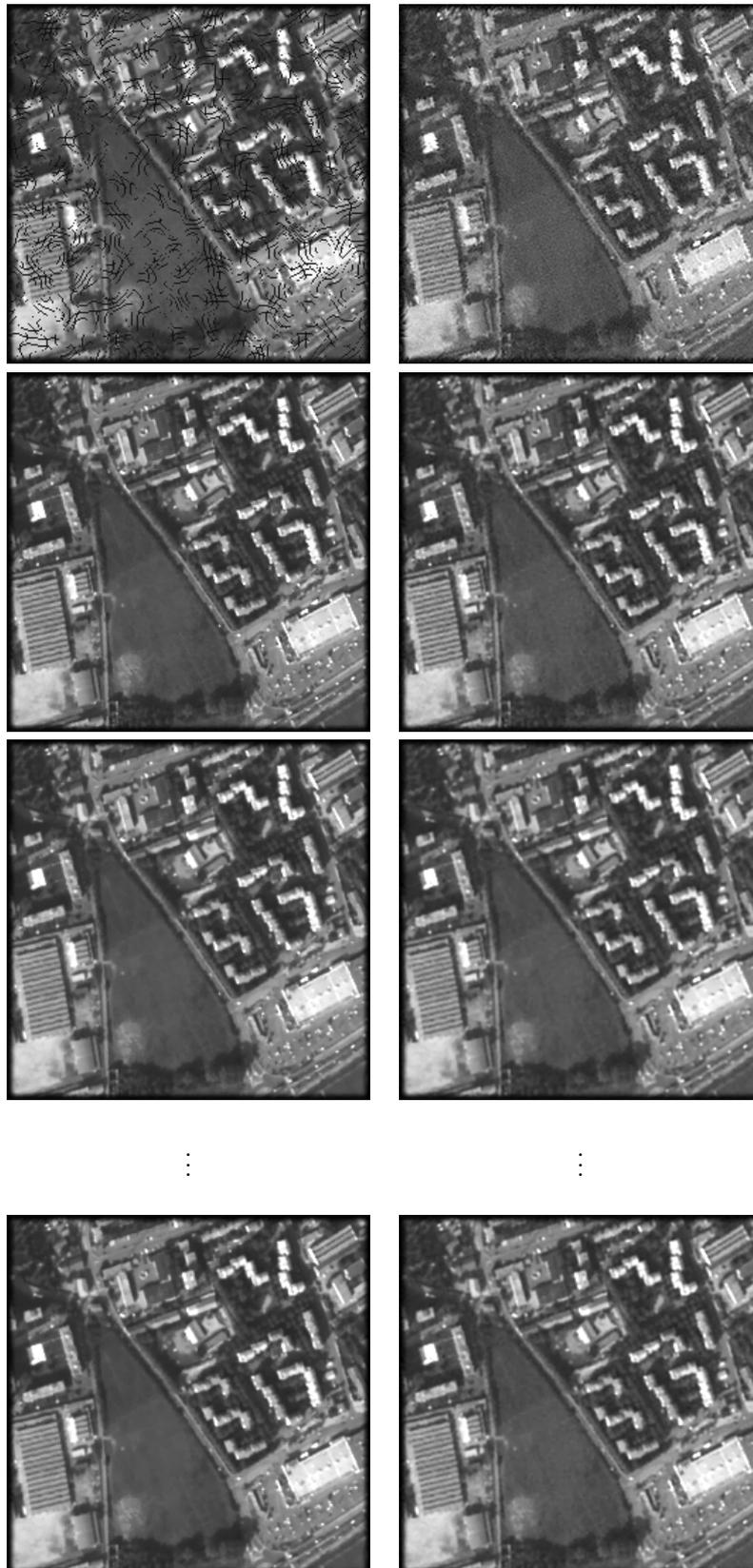


Figure 4.22: Results of our algorithm in the resampling (left) and deblurring+resampling case (right) with spline degree  $d_1 \in \{0, 1, 2, 7\}$  (from top to bottom). Results for  $d_1 = 0$  are clearly unsatisfactory and artifacts due to the sampling local density are visible. For  $d_1 \geq 2$  the results are very close and the pixelation effect present for  $d_1 \in \{0, 1\}$  nearly disappears when  $d_1 \geq 2$ .



Figure 4.23: Experiment with spline degree zero (no deblurring) : gaps in the sampling set together with the conjugate gradient produce initial guess (left) with holes. If we force the Lagrange parameter the TV regularization fills these holes. Here we display an intermediate result (middle) where the holes are partially filled and a full convergence result (right) in which the holes have completely disappeared (remaining artifacts are due to the pixelation effect).

Experiments show that the computation time has a low sensitivity to the spline order, although it slightly increases for orders greater than 3. We did not test the influence of the projection order  $d_2$  but it should have a similar effect on the quality of the restoration because of the convergence of interpolating splines in periodic spaces  $S_{per}^d(n, m)$  to the Dirichlet kernel.

Concerning the smoothing parameter, a new algorithm such as Nesterov's algorithm [69] would be necessary since the function  $TV_\mu$  is strongly convex and admits faster minimization schemes (with linear convergence rate). We denoted an improvement in statistics and a reduction of the stair-casing effect in resampling experiments for  $\mu \in ]0, 1]$ , the improvement was less significant in the deblurring case. A sharp estimation of this regularization parameter for a class of images would be necessary to add this feature as an automatically tuned parameter in our algorithm.

## 4.2.4 Perspectives

### 4.2.4.1 Total variation at half-integers

TV is usually implemented by finite differences on the integer grid, but there exist other versions such as the half-integer implementation of L. Moisan [67] which reduces the stair-casing effect. We use bicubic splines here together with this modified TV and study the influence of this variation on the performances of our algorithm.

$$TV_{\mathbb{Z}/2}(u) = \sum_{k=0}^{2n-1} \sum_{l=0}^{2m-1} \left| \nabla u \left( \frac{k}{2}, \frac{l}{2} \right) \right| \quad (4.42)$$

where the gradient can be computed either by finite differences on the half-integer grid, or exactly as a function of the spline coefficients by formula (2.19). The Lagrangian formulation of the sampling problem with TV penalization rewrites as

$$u^* = \underset{u \in S_{per}^3(n, m)}{\operatorname{argmin}} \quad \frac{1}{2} \|u(x_{k,l}, y_{k,l}) - I_{k,l}\|^2 + \lambda \int_{\Omega} |\nabla u(x, y)| dx dy \quad (4.43)$$

The norm of the gradient is not differentiable this is why we replaced it here by  $\sqrt{|\nabla u|^2 + C}$  ( $C > 0$  a very small constant) and obtain the following expression for the gradient of  $E_{TV}$  :

$$\frac{\partial}{\partial a_{k,l}} E_{TV}(a) = (S_\Lambda^t (S_\Lambda a - I))_{k,l} + \lambda \iint_\Omega \frac{\langle \nabla u(x,y), \nabla \beta_{3,2}(\cdot - (k,l)) \rangle}{\sqrt{|\nabla u(x,y)|^2 + C}} \quad \forall k,l \quad (4.44)$$

where  $u \in S_{per}^3(n, m)$  has spline coefficients  $(a_{k,l})$ . The computation of this integral is not elementary even in a spline space, where the exact expression of  $u$  is known. This is why we used the modified TV of L. Moisan on the half integer grid as a discretization of this integral with Riemann series :

$$TV_{\mathbb{Z}/2}(u) = \frac{1}{4} \sum_{q=0}^{2n-1} \sum_{p=0}^{2m-1} \sqrt{|\nabla u\left(\frac{q}{2}, \frac{p}{2}\right)|^2 + C} \quad (4.45)$$

$$\frac{\partial}{\partial a_{k,l}} E_{TV_{\mathbb{Z}/2}}(a) = (S_\Lambda^t (S_\Lambda a - I))_{k,l} + \frac{\lambda}{4} \sum_{q=0}^{2n-1} \sum_{p=0}^{2m-1} \frac{\langle \nabla u\left(\frac{q}{2}, \frac{p}{2}\right), \nabla \beta_{3,2}\left(\frac{q}{2} - k, \frac{p}{2} - l\right) \rangle}{\sqrt{|\nabla u\left(\frac{q}{2}, \frac{p}{2}\right)|^2 + C}} \quad \forall k,l \quad (4.46)$$

The computation of  $\nabla u\left(\frac{q}{2}, \frac{p}{2}\right)$  is a simple zoom operation on the gradient of  $u$ , it can be performed by convolution of the up-sampled spline coefficients as described in [89]. The gradient of  $\nabla \beta_{3,2}$  is computed thanks to separable filters and every convolution/scalar product uses periodic boundary conditions as we work in  $S_{per}^3(n, m)$ . The computation cost of the TV gradient can be compared to the computation of the gradient of the quadratic term

$$S_\Lambda^t (S_\Lambda a - I) \asymp 32 \times m \times n \quad \nabla E_{TV_{\mathbb{Z}/2}} \asymp (300 + 4\alpha) \times m \times n \quad (4.47)$$

with  $\alpha$  the cost of a square root computation (about 10 floating point operations). This modified TV has a great impact on the numerical complexity, here the gradient of the classical TV would have a computational cost similar to the gradient of the quadratic term.

We implemented the gradient descent algorithm 4.24 below, although faster algorithms exist for this strictly convex function. This gives an insight of the difference between  $TV$  and  $TV_{\mathbb{Z}/2}$  despite the very small smoothing constant  $C = 1.10^{-15}$  aimed at avoiding a division by zero. This smoothing is in fact not necessary if we use a sub-gradient descent algorithm or a dual method, which we had not heard of at the time of this implementation.

- Let  $a_0 \in \mathbb{R}^{n \times m}$
- For all  $n \geq 0$ 
  - $\rho_n = \frac{\alpha}{10+n}$
  - $a_{n+1} = a_n - \rho_n \nabla E_{TV_{\mathbb{Z}/2}}(a_n)$

Figure 4.24: Simple gradient descent with step size  $\rho_n \rightarrow 0$  following Polyak's rule.

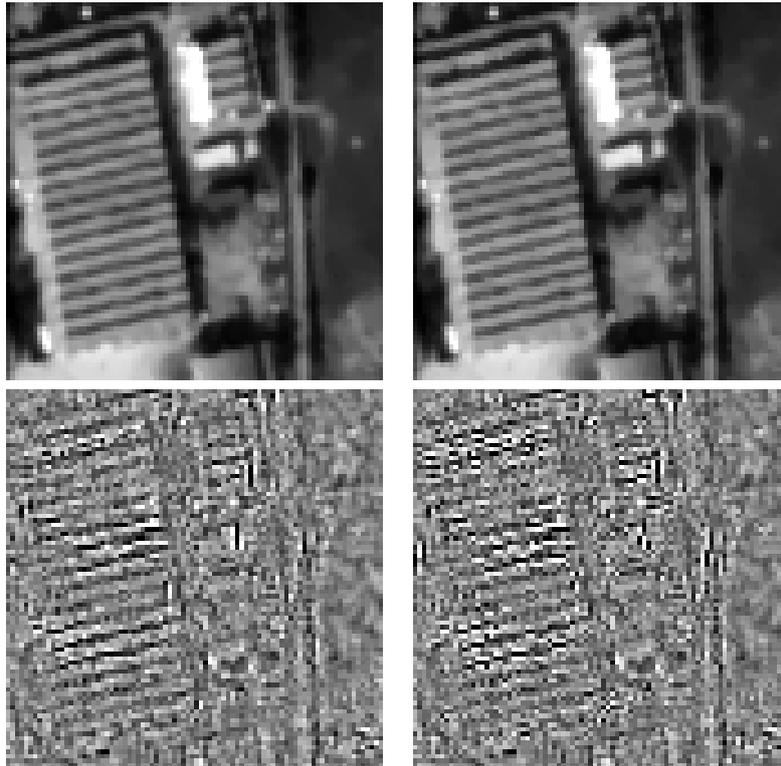


Figure 4.25: Results from the half integer TV (left) compared to classical TV (right) for low noise level ( $\sigma = 1$ ). The error term shows the advantage of half integer TV in the reconstruction of textured areas, there is almost no difference in smooth regions for this noise level. For higher noise levels the stair-casing effect is slightly reduced in smooth areas.

noise level	PSNR	
	$TV_{\mathbb{Z}/2}$	TV
1	<b>44.5339</b>	43.2283
3	<b>39.1153</b>	38.5255
5	<b>36.1033</b>	35.7112
7	<b>34.2309</b>	33.8419

Figure 4.26: Resampling with the TV at integers (left) and  $TV_{\mathbb{Z}/2}$  (right) for the same Lagrange parameters ( $10^4$  iterations in the gradient descent). Results displayed here use a very high number of iterations for both algorithms which emphasizes the stair-casing effect, this explains the differences between these PSNR results and results from table 4.2 but the phenomenon persists for lower convergence rates.

In our experiments, the half-integers TV always gave slightly better results than classical TV. The improvement should be more important in deblurring and zooming experiments as explained in [46]. An adaptation of our deblurring algorithm to this zoomed TV could be a new path of research but it is not as straight forward as the Quasi-Newton algorithm in [46]. Indeed the dual  $TV_{\mathbb{Z}/2}$  denoising problem in this case amounts to compute the following projection

$$S^* = \underset{S \in K_{TV}}{\operatorname{argmin}} \frac{1}{2} \|\lambda Z^* \nabla^* S - V\|^2$$

with  $S$  a  $2n \times 2m$  two-dimensional vector and  $Z$  the derivation+zooming operator on the considered spline space.

The conjugate derivation+zoom operator requires the computation of a convolution by means of FFT in periodic spline spaces, this would increase the cost of each iteration in the Projected Gradient algorithm in a non-negligible way. Other boundary conditions could prevent the use of Fourier transforms (symmetric or constant extension) but the impact on the computation time would still be significant 4.47, and the spline approach would be no more advantageous.

#### 4.2.4.2 Decomposition of the convolution operator

We examined the question of the computational cost of the gradient of the data-fitting term. Our formulation of the deblurring problem uses the Fourier transform for the diagonalization of the matrix  $H_{proj} M_{d_1}^{-1}$ , this implementation is certainly faster than the naive implementation because of the large support of the equivalent filter (given by a row of  $H_{proj} M_{d_1}^{-1}$  with periodic conditions) but can be very expensive for large image sizes.

Works of Malgouyres [64], Durand et al. [42], Kalifa et al. [54] have shown the efficiency of wavelet bases and wavelet packets [65] in the resolution of deblurring problems. The main idea is that wavelets have a quite good frequency localization and the blurring kernel can be decomposed as a series in the basis or family of wavelets.

We present here an application to functions in  $S_{per}^3(n, m)$  with  $n, m$  sufficiently large integers multiples of  $2^J$  and orthogonal wavelet bases (Daubechies wavelets). The idea is to approximate the projection in section 4.2.3.1 by an operator with very few coefficients in the wavelet domain thanks to the shape of the FTM. Indeed this FTM is smooth except at the origin which is exactly the part of the frequency domain refined by the classical wavelet transform (Mallat's cascade algorithm). Let  $f, g$  be the discrete filters associated to a multi-resolution analysis and  $a^{(0)}$  the signal to be decomposed, for simplicity we present the case of one dimensional signals with length  $N$  :

$$\begin{aligned} a^{(-1)}(l) &= (\tilde{f} * a^{(0)})(2l) & l = 0..N/2 - 1 \\ b^{(-1)}(l) &= (\tilde{g} * a^{(0)})(2l) & l = 0..N/2 - 1 \\ & \dots \\ a^{(-J)}(l) &= (\tilde{f} * a^{(-j+1)})(2l) & l = 0..N/2^j - 1 \\ b^{(-J)}(l) &= (\tilde{g} * a^{(-j+1)})(2l) & l = 0..N/2^j - 1 \end{aligned}$$

The reconstruction of  $a^{(0)}$  from  $a^{(-J)}$  and the  $\{b^{(-j)}\}_{j=1..J}$  uses the two properties of  $\hat{f}, \hat{g}$  : partition of unity and orthogonality [65] :

$$\begin{aligned} a^{(-j+1)}(l) &= \sum_{k=0}^{N/2^j-1} f(l-2k) a^{(-j)}(k) + \sum_{k=0}^{N/2^j-1} g(l-2k) b^{(-j)}(k) \\ & \dots \\ a^{(0)}(l) &= \sum_{k=0}^{N/2-1} f(l-2k) a^{(-1)}(k) + \sum_{k=0}^{N/2-1} g(l-2k) b^{(-1)}(k) \end{aligned}$$

We showed in section 4.2.3.1 that the projection of  $h * u$  rewrites as a convolution on the spline coefficients :

$$\text{proj}(h * u) = \sum_{k,l} (h_{proj} * a)[k,l] \beta_{d_2,2}(\cdot - (k,l))$$

If the  $a^{(0)}$  vector represents the spline coefficients of  $u$ , we propose to approximate the projection  $\text{proj}(h * u)$  with the following series :

$$\begin{aligned} \text{proj}(h * u)(x) &\approx \sum_j \sum_{l'} \left( \sum_k b^{(-j)}(l' - k) h^j(k) \right) \tilde{\psi}_{l'}^{(-j)}(x) + \sum_{l'} \left( \sum_k a^{(-j)}(l' - k) h^{j+1}(k) \right) \tilde{\phi}_{l'}^{(-j)}(x) \\ &\approx \sum_j \sum_{l'} (h^j * b^{(-j)})(l') \tilde{\psi}_{l'}^{(-j)}(x) + \sum_{l'} (h^{j+1} * a^{(-j)})(l') \tilde{\phi}_{l'}^{(-j)}(x) \end{aligned} \quad (4.48)$$

where  $h^j, h^{j+1}$  are discrete filters acting on the wavelet coefficients and  $\tilde{\psi}^{(-j)}, \tilde{\phi}^{(-j)}$  correspond to the convolution of  $h$  with of some continuous functions  $\psi^{(-j)}$  and  $\phi^{(-j)}$  defined by :

$$\begin{aligned} \phi_k^{(0)}(x) &= \beta^{(3)}(x - k) \quad k = 0..N - 1 \\ \phi_k^{(-1)}(x) &= \sum_l f(l) \phi_l^{(0)}(x - 2.k) \\ \psi_k^{(-1)}(x) &= \sum_l g(l) \phi_l^{(0)}(x - 2.k) \\ &\dots \\ \phi_k^{(-j)}(x) &= \sum_l f(l) \phi_l^{(-j+1)}(x - 2^j.k) \\ \psi_k^{(-j)}(x) &= \sum_l g(l) \phi_l^{(-j+1)}(x - 2^j.k) \end{aligned}$$

The approximation formula 4.48 is generally not verified but it is indeed verified for the Shannon multi-resolution analysis. This approximation of a blurring kernel in Shannon's wavelet basis amounts to approximate  $\hat{h}$  with trigonometric polynomials  $\hat{h}^j$  on each set  $[-\frac{\pi}{2^{j-1}}, -\frac{\pi}{2^j}] \cup [\frac{\pi}{2^j}, \frac{\pi}{2^{j-1}}]$ ,  $1 \leq j \leq J$  and  $[-\frac{\pi}{2^j}, -\frac{\pi}{2^j}]$ .

Filters  $h^j$  with chosen support size can be computed exactly thanks to the cascade algorithm and the same initial scalar products as in section 4.2.3.1. Our attempt to use small support filters  $h^j$  ( $3 \times 3$  and  $5 \times 5$ ) together with Daubechies wavelets revealed unfortunately quite bad with typical wavelet filter sizes (we tried filter sizes from 8 to 16). The aliasing effects nearly disappeared for large wavelet filters sizes but the method was no more competitive compared to the Fourier transform and has not been chosen in the final implementation. Experiments have shown that this accounts for only a small part of computations and does not prevent the algorithm from being much faster than the FAR method.

In fact every nonlinear process on the wavelet coefficients suffers from this aliasing effect since the orthogonality relation between wavelets is no more verified as pointed out in previous works from F. Malgouyres with simpler  $1 \times 1$  filters  $h^j$  (equivalent to a multiplication by a scalar). A better solution for further investigation could be the design of especially adapted wavelet filters (not necessarily orthogonal wavelets) or a different computation of filters  $h^j$  which would reduce the aliasing phenomenon compared to the classical  $L^2$  projection.

### 4.3 Conclusion

This study of the irregular sampling problems shows that splines are numerically interesting even in the deblurring case. The convolution operator can be handled with an approximation (here a projection) and the use of spatial regularization (quadratic penalization on the Hessian as in [5] and TV in our works) gives very good results. Problems  $\mathcal{P}_{\text{samp}}$  and  $\mathcal{P}_{\text{blur}}$  should not be treated the same way. Indeed we observed as in [46] that quadratic penalization gives better results than non-quadratic TV regularization when no deblurring is required. In the deblurring case we proposed a semi-implicite algorithm for the resolution of the subdifferential inclusion with automatically tuned parameters (Lipschitz constant, Uzawa's loop), together with a dual approach in the proximal computation. Many new results have appeared in the literature on these semi-implicit schemes and dual approaches since our work was achieved, for example [24] [8] [12]. Our algorithm could easily be adapted to these new schemes and provide an even faster implementation of the spline restoration.

In the perspectives we mention a variant of TV and show how it would increase the computational cost of our algorithm. We think that splines would no more be advantageous with this variant, trigonometric polynomials in this case would certainly have a comparable numerical complexity and allow for more general regularization functions (such as FAR). Finally, we observed slight differences in the behaviour of locally and globally constrained methods, this phenomenon could be due to the low sampling density and further investigation is necessary if a local approach with splines is examined.

## Chapter 5

# A study of bandpass signals in interferometry

### 5.1 An application of bandpass signals sampling in Static Interferometry

This last chapter of my thesis report is a summary of a R&T work in collaboration with D. Jouglet and C. Pierrangelo from CNES. The foundation of this R&T is the SIFTI project (Static Interferometry Fourier Transform Instrument) of a sounding device for the monitoring of  $CO$  and  $O_3$  concentration in the high atmosphere [16] [20] [51] [74].

The nature of the signals and the acquisition modes raise both theoretical and numerical questions, in particular the feasibility and stability of the reconstruction for the different sampling modes. The first part of this chapter is a short description of the physical phenomenon (interferometry), the SIFTI instrument and some of the questions that were brought to our attention. In a second part we provide answers to these questions in the form of a self-consistent article-like section, with several theoretical results on irregular sampling in bandpass spaces.

#### 5.1.1 Presentation of the SIFTI static interferometer

CNES<sup>1</sup> studied the feasibility of the SIFTI optical instrument (Static Infrared Fourier Transform Spectrometer) between 2007 and 2010 (Hébert et al. [51]). The scope of SIFTI is to monitor atmospheric pollution from space. SIFTI is a new kind of Fourier Transform Spectrometer (FTS) in which all mobile parts have been replaced by a 100% static concept. This concept imposes instrumental constraints on the measured signal and requires new processing which lead us to the present study. The SIFTI mission aims at providing an atmospheric sounding from space of ozone and carbon monoxide, in several places in the world with a revisit of a few hours. Such soundings may be performed thanks to spectra with high spectral and radiometric resolution in the thermal infrared, from which the tropospheric mixing ratios may be retrieved [77]. These spectra are given in radiance ( $S$ , in  $W.m^{-2}.sr^{-1}.cm$ ) or in brightness temperature (in  $K$ ) with respect to the sampled wavenumber ( $\sigma$  in  $cm^{-1}$ ). In [74], Pierangelo et al. expose the mission specification expressed by scientific users. The two gases are studied through two separate and narrow spectral bands in which each gas presents significant signatures : B1, dedicated to ozone, covers the  $1030 - 1070 cm^{-1}$  range (about  $9.5 \mu m$ ); and B2, dedicated to carbon monoxide, covers

---

<sup>1</sup>French Space Agency

the  $2140 - 2180 \text{ cm}^{-1}$  range (about  $4.6 \mu\text{m}$ ). The spectra of gases present very narrow and regular absorption or emission peaks as shown by figure 5.1. No assumption is made in the present study on the signal we want to retrieve, except for continuity and differentiability. This choice aims at avoiding any removal of unexpected feature. SIFTI is specified to a spectral apodized resolution of  $0.125 \text{ cm}^{-1}$  ( $0.0625 \text{ cm}^{-1}$  non apodized), with a spectral sampling of  $0.0625 \text{ cm}^{-1}$ . The radiometric requirements are  $\sim 0.1 \text{ K}$  (noise equivalent difference temperature) in B1 and  $\sim 0.2 \text{ K}$  NeDT in B2.

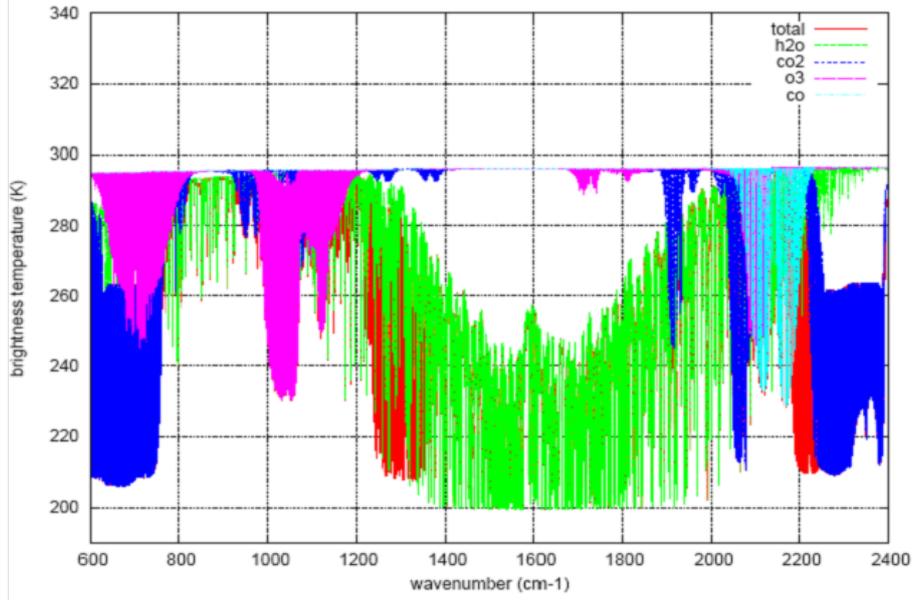


Figure 5.1: Example of atmospheric spectra for several elements among which  $CO$  and  $O_3$ .

Fourier Transform Spectrometry is based on the Michelson interferometer. The incident light flux is divided into two paths, or *arms*, by a separating plate. The light in the two arms encounters different optical path lengths before being recombined and then focused on a square detector as shown in figure 5.3 and illustrated by the famous Young's double slit experiment below.

These interferences can be expressed as a function of the oscillating electromagnetic fields corresponding to each arm. In the simplified case of a single frequency light source with (time) frequency  $f$  (in  $s^{-1}$ ), each arm corresponds to an electromagnetic field with amplitudes  $E_1(t)$  and  $E_2(t)$  at the sensor's location :

$$E_1(t) = e_1 \sin(2\pi f t + \phi_1) \quad E_2(t) = e_2 \sin(2\pi f t + \phi_2)$$

where the phase difference  $\phi_2 - \phi_1$  is linked to the OPD by the relation  $\phi_2 - \phi_1 = 2\pi \frac{OPD}{\lambda}$  with  $\lambda$  the wavelength ( $\lambda = \frac{c}{f}$  with  $c$  the light velocity).

The signal retrieved by the sensor is proportional to the energy of  $E_1 + E_2$  computed on a

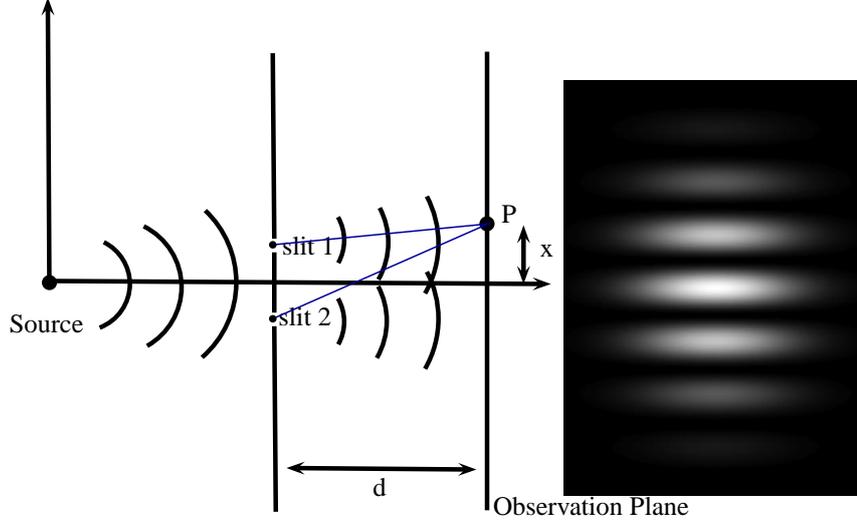


Figure 5.2: Interferences of light passing through two slits (Young's experiment 1801). This famous experiment illustrates the constructive (bright areas) and destructive interferences (dark areas) between two coherent light sources. The intensity of interferences is a function of the optical paths difference, the distance from the considered point  $P$  of the observation plane to each slit.

period  $T = \frac{1}{f}$  which produces the interference patterns.

$$\begin{aligned}
 I(OPD) &\propto \int_0^T (E_1(t) + E_2(t))^2 dt \\
 &\propto \int_0^T (e_1 \sin(2\pi f t + \phi_1) + e_2 \sin(2\pi f t + \phi_2))^2 dt \\
 &\propto \frac{T}{2} \left( e_1^2 + e_2^2 + 2 e_1 e_2 \cos\left(2\pi \frac{OPD}{\lambda}\right) \right)
 \end{aligned}$$

In the case of a polychromatic signal with continuous spectral distribution  $S(\sigma)$  ( $\sigma$  is the spatial frequency in  $cm^{-1}$  or wavenumber), the SIFTI sensor measures a collection of interferences with spatial frequencies belonging to a certain interval  $[\sigma_1, \sigma_2]$  called the *interferogram*  $I$  acquired for different OPDs.

$$I(OPD) = \int_{\sigma_{min}}^{\sigma_{max}} S(\sigma) (1 + \cos(2\pi\sigma OPD)) d\sigma \quad (5.1)$$

with  $I$  in  $W.m^{-2}.sr^{-1}$  and  $S$  in  $W.m^{-2}.sr^{-1}.cm$ .

This formula shows that the measured interferogram is the cosine transform of the searched spectrum. After conversion to complex values, the spectrum can be retrieved by the inverse Fourier transform of the interferogram. In classical Michelson interferometers like IASI (Blumstein [15], Simeoni [84]), the different OPDs between the two arms are obtained by a moving mirror in one or both arms (figure 5.3). The sampling of the interferogram is obtained for several positions of the mirrors at different successive times. In static interferometers like SIFTI, the moving mirror is replaced by a pair of crossed staircase fixed mirrors (figure 5.3). A laboratory breadboard proved the feasibility of this concept (Brachet et al. [16]). This evolution reduces parasitic dynamic

perturbations and increases the reliability of the instrument. The crossed staircase mirrors are imaged on a detector matrix on which several facets appear, each facet being the crossing of two steps, one in each arm of the interferometer. Each facet corresponds therefore to the value of the interferogram for a single OPD. The entire interferogram is acquired at once from the reading of the value of each OPD facet. The OPDs are obtained for B1 and B2 by a unique interferometer core, after which dichroic filters separates the two bands on two different detectors.

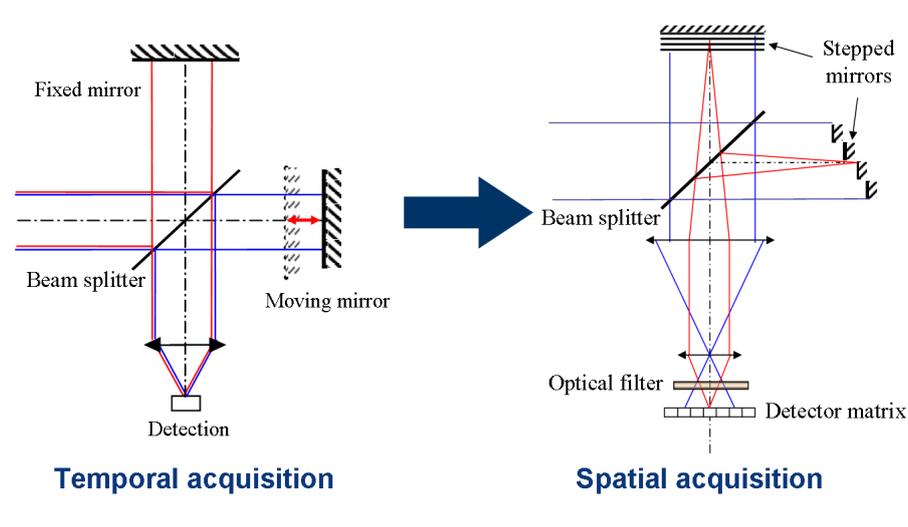


Figure 5.3: Michelson's static and dynamic interferometers

In every FTS, the spectral non apodized resolution  $\Delta\sigma$  imposes the value of the maximum OPD :  $\frac{1}{2\Delta\sigma}$ . An optical apodization reduces the contrast at high OPDs because of the wide field angle of SIFTI, but we do not to consider it for simplification. The numerical value for maximum OPD is 8 cm, which constrains the physical depth of the staircase mirrors. Static FTS imposes additional constrains to interferograms that have to be taken into account for the spectrum computation. First, the staircase mirrors have a limited number of steps ( $\sim 30$ ) due to manufacturing design. The OPDs obtained by the crossing of both mirrors are therefore limited to a number of  $\sim 1000$ . As a comparison, IASI provides more than 50,000 samples in the interferogram. The SIFTI mission needs for high spectral resolution and narrow spectral bands can be well accommodated with this small number of samples by 1) acquiring only single-sided interferograms in the positive OPDs, and by 2) applying the generalized Shannon theorem. According to this theorem, the signal may be strongly under-sampled without any loss of information if

$$\Delta_{OPD} > \frac{1}{2(\sigma_{max} - \sigma_{min})} \quad (5.2)$$

$$\exists k \in \mathbb{N} \quad \frac{k}{2\Delta_{OPD}} < \sigma_{min} < \sigma_{max} < \frac{k+1}{2\Delta_{OPD}}$$

with  $\Delta_{OPD}$  the distance between two samples of the OPD. Under these conditions, several low frequency aliases of the useful signal appear but their small width prevents any overlapping. These conditions makes the concept of static FTS well suited to the SIFTI mission. In SIFTI,  $\Delta_{OPD} = 80 \mu m$  defines the Nyquist frequency  $\sigma_{nyquist}$  equal to  $62.5 cm^{-1}$  while  $[\sigma_{min}, \sigma_{max}] \sim [1030, 1070] cm^{-1}$ . A second constraint of the static FTS appears in the design of the staircase mirrors which is of limited precision. The positions of the steps are slightly irregular with respect

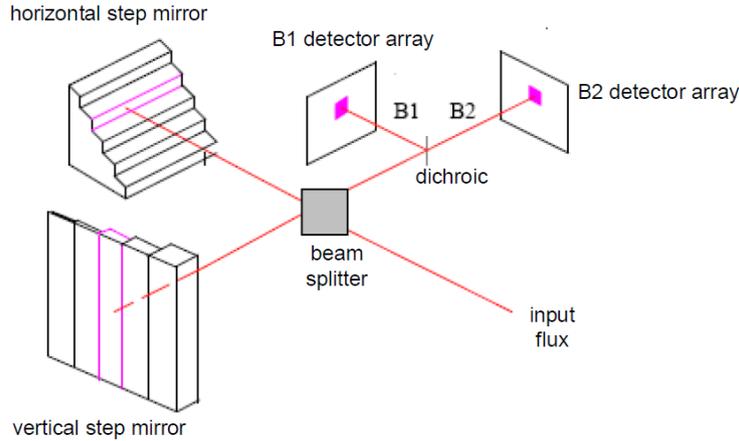


Figure 5.4: The static interferometry instrument concept uses fixed stepped mirrors instead of a moving plane mirror. The incoming light is filtered and separated into two arms which are then reflected by the stepped mirrors and the beam splitter in direction of the detection array. The two mirrors have different step sizes, one of them has small steps with height  $dx/2$  and the other one has larger steps, each large step is equal to the sum of all small steps. ( figure from [16] )

to the requested regular positions, which results in a slightly irregular OPD sampling of the interferogram. In our preliminary studies, we proved that the OPDs can be modeled at each sample by a Gaussian law, whose mean is the regular expected OPD and whose standard deviation is  $5\ \mu\text{m}$ . The irregular OPD values will be known accurately (up to  $10\ \text{nm}$ ) in the flight instrument.

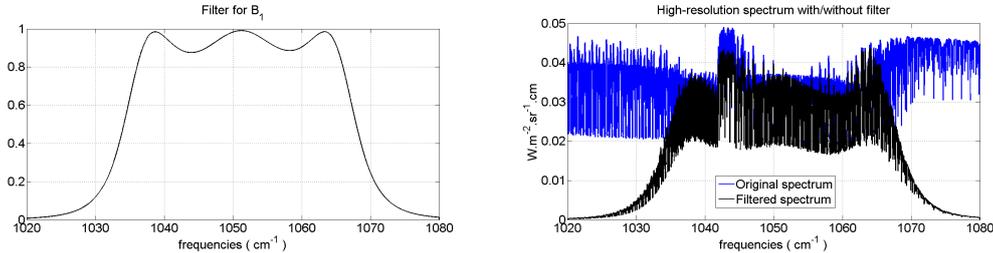


Figure 5.5: Multiplicative filter on the spectrum (top) and the high-resolution spectrum (bottom). This filtering takes place before the intensity measurements by the matrix detector and is thus taken into account in the simulation of the interferogram, together with sampling perturbations, noise and possible apodization.

A sampling improvement has been proposed by Cansot et al. [20] which consists in the concatenation of two identical sampling combs shifted by  $\lambda/4$  with  $\lambda = 1/\sigma_{mid} = 2/(\sigma_{max} + \sigma_{min})$  (phase opposition), as illustrated by figure 5.6. Each comb presents the same characteristics of under-sampling and slight irregularities as described in the previous paragraphs, but may be sparser. Empirical studies at CNES proved that using this sampling reduces the radiometric noise amplification from interferogram to spectrum. This variant is called IIS (*Interferogram Interlaced Sampling*), while the nominal one is called IES (*Interferogram Even Sampling*). Such a sampling is obtained in practice by the phase modulation created by an optical coating deposited on the half of each step of the crossed mirror, splitting it into steps of same length but half widths. Each facet is split into four values, which are linearly combined to provide the expected doublets.

$$\Lambda_{IES} = \{x_k\} \quad \text{with} \quad x_k = x_{min} + k dx + \nu_k \quad 0 \leq k \leq N_s - 1 \quad x_{min} \in [-0.13, 0] \text{ cm} \quad (5.3)$$

The second sampling mode, called Interferogram Interlaced Sampling (IIS), consists in two shifted grids with inter-doublet step  $dx$  and intra-doublet step  $\epsilon$  :

$$\Lambda_{IIS} = \{x'_k\} \quad \text{with} \quad x'_{2k} = x_{min} + k dx + \nu_k \quad x'_{2k+1} = x_{2k} + \epsilon \quad (5.4)$$

$$0 \leq k \leq \frac{N_s}{2} - 1$$

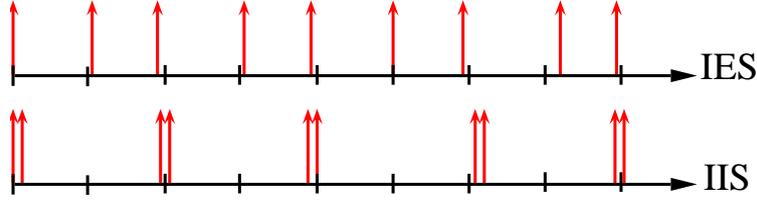


Figure 5.6: The two main sampling schemes we study here i) perturbed regular sampling (IES) described in eq. (5.3) and ii) perturbed interlaced sampling or nonuniform sampling (IIS) described in eq. (5.4) where the consecutive samples  $x'_{2k}$  and  $x'_{2k+1}$  are always separated by the same gap  $\epsilon > 0$ .

To summarize, the SIFTI instrument amounts to compute a spectrum through the Fourier Transform of an interferogram irregularly sampled and acquired in the Shannon's generalized theorem conditions. The irregularity of the sampling makes the FFT algorithm useless without prior processing and, associated with the undersampling, it prevents the numerical computation of Fourier transforms as integrals. SIFTI therefore rises a new challenge in signal processing that this paper contributes to solve.

### 5.1.2 CNES's state of the art

Two methods have been proposed at CNES for the retrieval of spectra in such conditions. The first method computes the interferogram on a regular sampling from the measurements and then applies an FFT to get the spectrum. The regularization is based on the interpolation formula from a regular sampling, which gives the interferogram value  $I$  at any OPD from the convolution between the  $I_{reg}$  values at the regular positions  $OPD_{reg}$

$$I(OPD) = \sum_k I(OPD_{reg}(k)) \Phi(OPD - OPD_{reg}) \quad (5.5)$$

with  $\Phi(\cdot) = \text{sinc}(\pi(\sigma_{max} - \sigma_{min}) \cdot) \cos(2\pi \sigma_{min} \cdot)$ . We express this formula at the instrumental irregular OPDs  $\{OPD_{irreg}\}$  and compute the unknown vector  $I_{reg}$  by pseudo-inversion :

$$I_{irreg} = Q I_{reg} \quad (5.6)$$

with  $Q_{k,l} = \Phi(OPD_{irreg}(k) - OPD_{reg}(l))$ . Then we take

$$\begin{aligned} I_{reg} &= Q^\dagger I_{irreg} \\ &= (Q^t Q)^{-1} M^t I_{reg} \end{aligned} \quad (5.7)$$

if  $(Q^t Q)$  is invertible. The quality of the result depends on the condition number of  $Q Q^t$ . To increase this conditioning  $I_{irreg}$  has more values than  $I_{reg}$ .

The second method avoids any FFT and considers the relationship between  $S_{reg}$ , the vector of the regular samples of the spectrum at  $\{\sigma_j\}$  with  $\sigma_j = \sigma_{min} + j\Delta\sigma$ , and  $I_{irreg}$  the vector of the irregular samples of the interferogram at  $\{OPD_{irreg}(k)\}$ . The cosine transform of the spectrum can be calculated as numerical integrals because  $\{\sigma_j\}$  is regular. A matrix relationship can therefore be derived

$$I_{irreg} = R S_{reg} \quad (5.8)$$

where  $R$  is defined by  $R_{i,j} = \cos(2\pi OPD_{irreg}(i) \sigma_j) \Delta\sigma$ . Obtaining the regular samples of the spectrum from the irregular samples of the interferogram is an inverse problem that can be solved by optimal estimation

$$\begin{aligned} S_{reg} &= R^\dagger I_{irreg} \\ &= R^t (R R^t)^{-1} I_{irreg} \end{aligned} \quad (5.9)$$

if  $(R R^t)$  is invertible. Here we consider that the matrix  $R$  has more columns than rows (the discrete spectrum has more points than the irregular interferogram). Once again the quality of the result depends on the conditioning of the  $(R R^t)$  matrix.

### 5.1.3 Questions

The early implementation of these two methods in the SIFTI sampling conditions has raised both theoretical and numerical problems. Because of the irregularity in the design of the mirrors steps the reconstruction can be unstable and one can easily imagine that a strong irregularity may produce non-invertible systems. At this time the conditions for a feasible reconstruction are unknown for the kind of signals we are dealing with.

The second problem is the evaluation of each method : experiments on a simulated spectrum with very high resolution have shown that the two methods produce very similar results but have different sensitivities to noise in the interferogram and to irregularity in the OPDs. Because of the mission specifications we need sharp error estimations for each method and we would like to understand why the two reconstructions are generally close but different.

And last we want to understand what are precisely the sampling conditions in the IIS case and the theoretical justification for the good performances of this sampling relatively to noise and irregularity.

### 5.1.4 Contents of this chapter

Answering these questions demands a rigorous mathematical modelization of the signals and measurements in all the cases we presented, which are regular and irregular IES and IIS samplings in the presence of noise. In the following we first introduce the notations and remind the important theorems in bandpass sampling like the famous generalized Shannon's theorem mentioned earlier, but also other results well known from the Signal Processing community.

Then we consider the existing CNES's methods and show that they belong to a more general framework (Nonharmonic Fourier Series) which benefits from many mathematical results in irregular sampling in the bandlimited case. We extend the Duffin-Shaeffer perturbation theorem [41] to the bandpass case and study from a numerical point of view the behaviour of some reconstruction operators relatively to the irregularity in the OPDs.

The last part of this chapter is an attempt to justify the observations concerning sensitivity to noise in the interferogram. Some modelizations seem to be numerically less sensitive to these parameters, and a new one from the general nonharmonic Fourier series framework is analysed.

## 5.2 Sampling in $BP(A, B)$

In the following we will note  $BP(A, B)$  the subset of  $L^2(\mathbb{R})$  of functions with Fourier transform supported in  $\Omega = [-B, -A] \cup [A, B]$ ,  $0 < A < B$ . Its Fourier transform  $\mathcal{F}(BP(A, B))$  is closed in  $L^2(\mathbb{R})$  and it is a Hilbert space for the scalar product on  $\Omega$

$$\langle f, g \rangle_{L^2(\Omega)} = \int_{\Omega} f(t) \overline{g(t)} dt.$$

### 5.2.1 Link with Static Interferometry

Consider the basic interferometry model for a signal  $S$  with support in  $[A, B]$  at locations  $\{x_k\}_{k=1..M}$

$$I_k = \int_A^B S(\sigma) \cos(2\pi x_k \sigma) d\sigma. \quad (5.10)$$

The spectrum energy  $S$  is measured by scalar products with a cosine function on  $[A, B]$ , which are in fact values of the Fourier transform of  $\tilde{S}$  the even part of  $S$  (with support in  $\Omega$ ) at irregular locations.

$$\begin{aligned} I_k &= \int_A^B S(\sigma) \left( \frac{e^{i2\pi x_k \sigma} + e^{-i2\pi x_k \sigma}}{2} \right) d\sigma \\ &= \frac{1}{2} \int_{\Omega} (S(\sigma) + S(-\sigma)) e^{i2\pi x_k \sigma} d\sigma \\ &= \mathcal{F}(\tilde{S})(x_k) := F(x_k). \end{aligned}$$

The fact that  $S$  is real-valued implies that the Fourier transform has Hermitian symmetry  $F(-x) = F(x)^*$ ,  $x \in \mathbb{R}$  and the knowledge of  $F$  at positive locations is indeed sufficient. We see that our problem is equivalent to recovering the Fourier transform  $\hat{f}$  of an even function  $f$  in  $BP(A, B)$  from the data  $\{f(x_k)\}_{k=1..N}$ .

### 5.2.2 Periodic Sampling

Functions whose Fourier transform has support in the union of intervals are called bandpass signals in Signal Processing (Kohlenberg [57], Vaughan [92], Vaidyanathan [62] et al.). Uniform and periodic nonuniform sampling of these signals correspond to nonperturbed IES and IIS sampling modes (see section 5.1) and are ruled by the two following theorems [92].

#### **Theorem 15** Generalized Shannon Theorem

Let  $0 < A < B$ ,  $dx > 0$ . Any function  $f$  in  $BP(A, B)$  can be reconstructed from its samples  $\{f(k dx)\}_{k \in \mathbb{Z}}$  if and only if  $dx$  belongs to one of the intervals

$$\left[ \frac{k}{2A}, \frac{k+1}{2B} \right], \quad 0 \leq k \leq \frac{A}{B-A}, \quad k \in \mathbb{N} \quad (5.11)$$

The sampling step  $dx$  may belong to the union of at most  $\lfloor \frac{A}{B-A} \rfloor$  intervals: these values guaranty that  $\Omega$  and all the shifted  $\Omega + \frac{k}{dx}$  ( $k \in \mathbb{Z} \setminus \{0\}$ ) do not overlap in Poisson's summation formula : this is a non-aliasing criterion.

We take the following example which includes the band B1 :  $[A, B] = [1020, 1080] \text{ cm}^{-1}$ . The generalized Shannon's theorem gives 17 intervals of the form  $[\frac{k}{2040}, \frac{k+1}{2160}]$ , and the last interval is reduced to the point  $\{\frac{17}{2040}\} = \{\frac{1}{120}\}$ .

Periodic nonuniform sampling rule is more complicated because it depends on the way the two intervals  $[-B, -A]$  and  $[A, B]$  overlap when  $\Omega$  is periodized with a step  $\frac{1}{dx}$ , as stated in the following theorem.

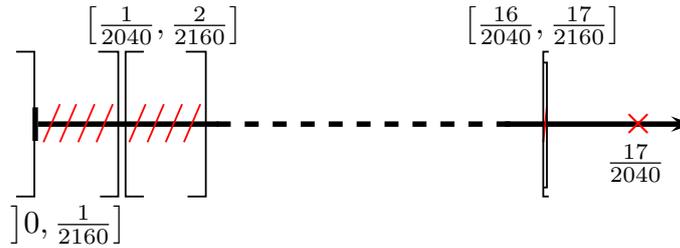


Figure 5.7: Example of Shannon's intervals with  $[A, B] = [1020, 1080] \text{ cm}^{-1}$ . The last point is  $83,3 \mu\text{m}$  and the last interval is approximately  $[78.4, 78, 7] \mu\text{m}$ .

**Theorem 16** *Periodic Nonuniform Sampling*

Let  $0 < A < B$ ,  $0 < dx \leq \frac{1}{B-A}$  and  $d > 0$ . Consider the set  $J = J(A, B, dx)$  of indices  $k$  for which the translation of step  $\frac{k}{dx}$  leads to non trivial overlapping.

$$J = \left\{ k \in \mathbb{Z} / \mu \left( \left( [-B, -A] + \frac{k}{dx} \right) \cap [A, B] \right) > 0 \right\}$$

with  $\mu$  the Lebesgue's measure. Then  $\Lambda = \{l dx\}_{l \in \mathbb{Z}} \cup \{l dx + d\}_{l \in \mathbb{Z}}$  allows reconstruction of any function  $f$  in  $BP(A, B)$  from its samples  $\{f(x_k)\}_{x_k \in \Lambda}$  if and only if

$$J \cap \frac{dx}{d} \mathbb{Z} = \emptyset$$

Two special cases are detailed in [92], integer and half-integer band positioning, which correspond to the case  $\frac{B+A}{B-A} \in \mathbb{N}$  and  $J = \{\frac{B+A}{B-A}\}$ . The authors computed the best  $d$  in this case in terms of inversion norm

$$d \in \frac{1}{4f_c} + \frac{1}{2f_c} \mathbb{Z} \quad f_c = \frac{B+A}{2} \quad (5.12)$$

and verified experimentally that this "quadrature" sampling leads to the best precision in the reconstruction. This result comes as a justification for the choice of intra-doublets spacing  $d = \frac{\lambda}{4}$  described in section 5.1.1 in the IIS case. This will be detailed in section 5.3.4.

In the general case of periodic nonuniform sampling (when  $\frac{B+A}{B-A} \notin \mathbb{N}$ ), the quadrature sampling is no more optimal and the best shift  $d$  realizes the minimal norm of an aliasing matrix :

$$\begin{aligned} & \inf_d \max_{k \in J} \left\| \begin{pmatrix} 1 & 1 \\ 1 & e^{-i2\pi d \frac{k}{dx}} \end{pmatrix}^{-1} \right\|_{2,2} \\ & = \inf_d \max_{k \in J} \frac{1}{2|\sin(\pi d \frac{k}{dx})|} \sqrt{2 + |2 \cos(\pi d \frac{k}{dx})|} \end{aligned} \quad (5.13)$$

with  $d$  such that  $J \cap \frac{dx}{d} \mathbb{Z} = \emptyset$ . Unfortunately there is no general formula for the best  $d$  in the case  $|J| > 1$ .

Reconstruction formulas are different in the uniform and nonuniform case. In the regular sampling case, a single reconstruction function  $\Phi$  is necessary but in the nonuniform case other functions  $\Phi_1$  and  $\Phi_2$  with piecewise constant Fourier transform are necessary to reconstruct the signal (see [92]). When conditions of theorems 15 or 16 are verified, any function  $f \in BP(A, B)$  is the limit of a uniformly converging series on every compact subset of  $\mathbb{R}$  :

$$f(\cdot) = \sum_{k \in \mathbb{Z}} f(k dx) \Phi(\cdot - k dx)$$

or

$$f(\cdot) = \sum_{k \in \mathbb{Z}} (f(k dx) \Phi_1(\cdot - k dx) + f(k dx + d) \Phi_2(\cdot - k dx - d)) \quad (5.14)$$

In [32] and [62] the authors generalized nonuniform sampling to multiband signals by considering an arbitrary union of regular sampling sets of the form

$$\Lambda = \bigcup_{1 \leq i \leq n} \{k dx + d_i\}_{k \in \mathbb{Z}} \quad (5.15)$$

and give a decomposition formula similar to (5.14) with  $n$  basis functions  $\Phi_i$ ,  $1 \leq i \leq n$ . The goal of this generalization is to approach the Landau's sampling density  $\mu(\Omega)$  which is the smallest density of a stable sampling set [58]. Note that the minimum density set of the form (5.15) always exists in the case of two bands with equal length, but not necessarily in the general multiband case [52].

### 5.2.3 Perturbed Sampling

A lot of works exist about perturbed sampling in bandlimited spaces, starting from Paley and Wiener [73], Duffin and Schaeffer [41], Kadec [53] and many others, to the more recent works of Favier and Zalik [47], Chui and Shi [27] who generalized the perturbation results in dimension one to higher dimensions. But only a few articles deal with irregular sampling in bandpass spaces, and more particularly perturbed sampling of the form

$$x_{k,i} = k dx + d_i + \xi_{k,i} \quad 1 \leq i \leq n \quad k \in \mathbb{Z} \quad (5.16)$$

with  $\xi_{k,i}$  i.i.d. random variables with compactly supported distribution. Martin et al. [94] have observed very interesting experimental results concerning such sampling sets. For large enough perturbations, the perturbed regular sampling set  $\{k dx + \xi_k\}$  allows reconstruction of a bandpass signal with the only restriction on  $\frac{1}{dx}$  to be slightly larger than  $\mu(\Omega)$ , whereas regular sampling must obey the generalized Shannon's condition (5.11).

But what happens if the sampling set already allows reconstruction? This is the problem of optimal bound on the perturbation. In dimension one the best result about perturbed sampling of bandlimited functions with maximum frequency  $B$  (noted  $BF(B)$ ) is Kadec's theorem [53].

**Theorem 17** *Kadec's 1/4 Theorem*

Let  $B > 0$ ,  $\Lambda = \{x_k\}_{k \in \mathbb{Z}}$  be a subset of  $\mathbb{R}$  such that

$$|x_k - \frac{k}{2B}| \leq L \quad \forall k \in \mathbb{Z}$$

If  $L < \frac{1}{8B}$  then the family  $\{e^{i2\pi x_k \cdot}\}_{k \in \mathbb{Z}}$  is a Riesz basis of  $L^2([-B, B])$  and there exists a Riesz basis  $\{g_k\}_{k \in \mathbb{Z}}$  of  $BF(B)$  such that

$$f(\cdot) = \sum_{k \in \mathbb{Z}} f(x_k) g_k(\cdot)$$

The bound  $\frac{1}{8B}$  is tight due to an example by Levinson (see [60] p. 65) and the name 1/4-theorem comes from the ratio  $\frac{1}{8B} / \frac{1}{2B} = 1/4$ . Levinson showed that if  $L = \frac{1}{8B}$  the sampling set captures all the signal information and Young [95] also showed that if  $L > \frac{1}{8B}$ , then there exist sampling sets that do not capture all the information in the signal. But we want to sample functions in  $BP(A, B)$  at a much lower density than  $2B$ , close to  $2(B - A)$ .

The  $\frac{\ln 2}{2\pi B}$ -perturbation theorem of Duffin and Schaeffer [41] can be adapted easily to the bandpass case since it only requires analytic functions with exponential growth (bandlimited functions) and a tight frame of  $L^2(\Omega)$ . It is well known [66] that the stable sampling problem in bandlimited spaces is equivalent to the frame property of a family of complex exponentials in the Fourier domain. We remind here the definitions of stable sampling set, frame and the frame property (see [58] and [66]).

**Definition 13** *Stable Sampling Set*

Let  $E$  be a subset of  $\mathcal{C}(\mathbb{R}) \cap L^2(\mathbb{R})$  and  $\Lambda = \{x_k\}_{k \in \mathbb{Z}}$  a subset of  $\mathbb{R}$ .  $\Lambda$  is called a stable sampling set of  $E$  if  $\sum_{k \in \mathbb{Z}} |f(x_k)|^2 < \infty$  for all  $f \in E$  and if there exists  $C > 0$  such that

$$\|f\|_{L^2(\mathbb{R})} \leq C \|f(x_k)\|_{l^2(\mathbb{Z})} \quad \forall f \in E$$

**Definition 14** *Frame*

Let  $\mathcal{H}$  be a  $\mathbb{C}$ -Hilbert space and  $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$  a subset of  $\mathcal{H}$ .  $\mathcal{F}$  is a frame of  $\mathcal{H}$  if there exist  $C, D \in \mathbb{R}$  with  $0 < C < D$ , such that

$$C \|h\|_{\mathcal{H}}^2 \leq \sum_{k \in \mathbb{Z}} |\langle h, f_k \rangle|^2 \leq D \|h\|_{\mathcal{H}}^2 \quad \forall h \in \mathcal{H}$$

**Theorem 18** *Frame Property*

Let  $\Lambda = \{x_k\}_{k \in \mathbb{Z}}$  be a subset of  $\mathbb{R}$ ,  $0 < A < B$  and  $\Omega = [-B, -A] \cup [A, B]$ .  $\Lambda$  is a stable sampling set of  $BP(A, B)$  if and only if  $\{e^{i2\pi x_k \cdot}\}_{k \in \mathbb{Z}}$  is a frame of  $L^2(\Omega)$ .

We can choose  $dx$  such that  $\{e^{ikdx \cdot}\}_{k \in \mathbb{Z}}$  forms a frame (in that case it is a tight frame) and get the following perturbation theorem.

**Theorem 19** *Perturbed IES Sampling*

Let  $0 < A < B$  and  $dx > 0$  satisfying the generalized Shannon's condition (5.11). The set  $\Lambda = \{x_k\}_{k \in \mathbb{Z}}$  such that

$$|x_k - k dx| \leq L \quad \forall k \in \mathbb{Z}$$

is a stable sampling set of  $BP(A, B)$  if  $L < \frac{\ln 2}{2\pi B}$ .

We refer the reader to appendix 5.6.1 for the proof of this theorem which uses Taylor series, as in [41].

At this time this is the only perturbation constant we know for bandpass spaces and the problem of the optimal perturbation bound for a given sub-Nyquist uniform sampling set is still not solved. The same problem occurs in periodic nonuniform sampling, where it would be useful to know the order of magnitude of the maximum perturbation. This case will be treated in details in section 5.3.4.

In our example  $[A, B] = [1020, 1080] \text{ cm}^{-1}$  with finitely many samples, the bound  $L = \frac{\ln 2}{2\pi B}$  is approximately  $1.02 \cdot 10^{-4} \text{ cm}$ , which is five times smaller than the standard deviation of the sampling location for the SIFTI interferometer. If we consider that no irregularity larger than 3.5 times the standard deviation occurs, we are very far from this bound (by a factor 17.5). Though, we are able to correctly reconstruct the signal (figures 5.8 and 5.9).

The observations of Martin et al. can be compared to the recent results of Meyer and Mattei [66], who have shown that there exist universal sampling sets called quasicrystals for the kind of functions we deal with. The main result of [66] is that any quasicrystal  $\Lambda$  with density  $\text{dens}(\Lambda)$  is a stable sampling set for functions in  $L^2(\mathbb{R}^n)$  with Fourier transform supported in a compact  $K$  provided that  $\text{dens}(\Lambda) > |K|$ .

In the bandpass case, experiments tend to favour perturbed sampling as a universal sampling method, especially when the perturbation is larger than the theoretical bound (here  $\frac{\ln(2)}{2\pi B}$ ). This perturbed set keeps its good properties when the original sampling set is stable, and becomes stable when the original set is not. This is not in contradiction with the results of Meyer and Mattei and makes easier the design of universal sampling sets.

We point now an important difference between regular and perturbed sampling : because of irregularity, there does not exist a single function  $\Phi_1$  such that the following reconstruction formula holds :

$$f(\cdot) \neq \sum_{k \in \mathbb{Z}} f(x_k) \Phi_1(\cdot - x_k)$$

but there exists a more general decomposition formula which uses analytic functions defined by infinite products of monomials  $\{(\cdot - x_k)\}$ . In the case of a finite number of samples, this corresponds to Lagrange's interpolation : this method is unfortunately too unstable. But in some cases (we give an example in section 5.4) it is possible to derive the expression of this decomposition from the infinite sampling set.

## 5.3 Reconstruction from a small number of irregular samples

### 5.3.1 A study of CNES methods

We consider here the two methods described in section 5.1.2. The CNES first method uses a linear system deriving from the truncated series (5.14), when  $dx$  satisfies the generalized Shannon's condition (5.11) .

$$\begin{aligned} (I_k) = (F(x_k)) &\approx \sum_{l=-N}^N F(l dx) \Phi(x_k - l dx) \\ &\approx Q \times (F(l dx))_{l=-N..N} \end{aligned}$$

where the matrix  $Q$  has general term  $Q_{k,l} = \Phi(x_k - l dx)$ . Since  $F$  is real and even, the system can be reformulated as

$$(I_k)_{k=1..M} \approx Q' (F(l dx))_{l=0..N} \quad (5.17)$$

with  $Q'_{k,l} = \Phi(x_k - l dx) + \Phi(x_k + l dx)$  if  $l \neq 0$  and  $Q'_{k,0} = \Phi(x_k)$ .

We get an approximation of the vector  $(F(l dx))_{l=0..N}$  by inversion of the system. The number of unknowns is usually equal to the number of observations ( $N = M$ ) , so the inversion can be performed :

$$(F(l dx)) = (Q')^{-1}(I_k)$$

but because of the redundant sampling or the bad condition number of the system it is better to take the pseudo-inverse of  $Q'$  with a chosen threshold.

The inverse Fourier transform of  $F$  can be observed on any output grid by the formula

$$\tilde{S}_{rec}(\cdot) = F(0) + 2 \sum_{l=1}^N F(l dx) \cos(2\pi l dx \cdot) \quad \text{on } \Omega \quad (5.18)$$

but it can also be observed on the base band by periodization with the fast Fourier transform (see figure 5.8(a) ).

The second CNES method for the recovering of  $\tilde{S}$  is an approximation of the integral in (5.10) by a Riemann integral

$$I_k \approx \sum_{0 \leq l < \lfloor \frac{B-A}{\Delta\sigma} \rfloor} \tilde{S}(A + l\Delta\sigma) \Delta\sigma \cos(2\pi x_k(A + l\Delta\sigma))$$

This approximation makes sense when  $\Delta\sigma$  is small, and depends on the sup norm of  $\tilde{S}'$ . We know that the number of samples  $M$  in the device is not sufficient (even in the nonperturbed case) to catch all the information in  $\tilde{S}$ . In the CNES experiments the discretization of  $\tilde{S}_{rec}$  between 1020 and 1080  $cm^{-1}$  had as many points as the number of samples in the interferogram, typically less than 1600 points. This is clearly not enough and leads to the resolution of the linear system

$$(I_k)_{k=1..M} \approx Q'' (\tilde{S}_{rec}(A + l\Delta\sigma))_{0 \leq l < \lfloor \frac{B-A}{\Delta\sigma} \rfloor} \quad (5.19)$$

with  $Q''_{k,l} = \Delta\sigma \cos(x_k(A + l\Delta\sigma))$ . The matrix  $Q''$  should be rectangular with much more columns than rows to get a good approximation of the Riemann integral (our experiments showed that a factor 75 was enough) and must be inverted by pseudo-inversion (two reconstructions of the same interferogram are displayed in figure 5.9, both methods have 976 unknowns, which corresponds to the critical sampling  $dx = 83,3 \mu m$  between  $[0, 8] cm$ ).

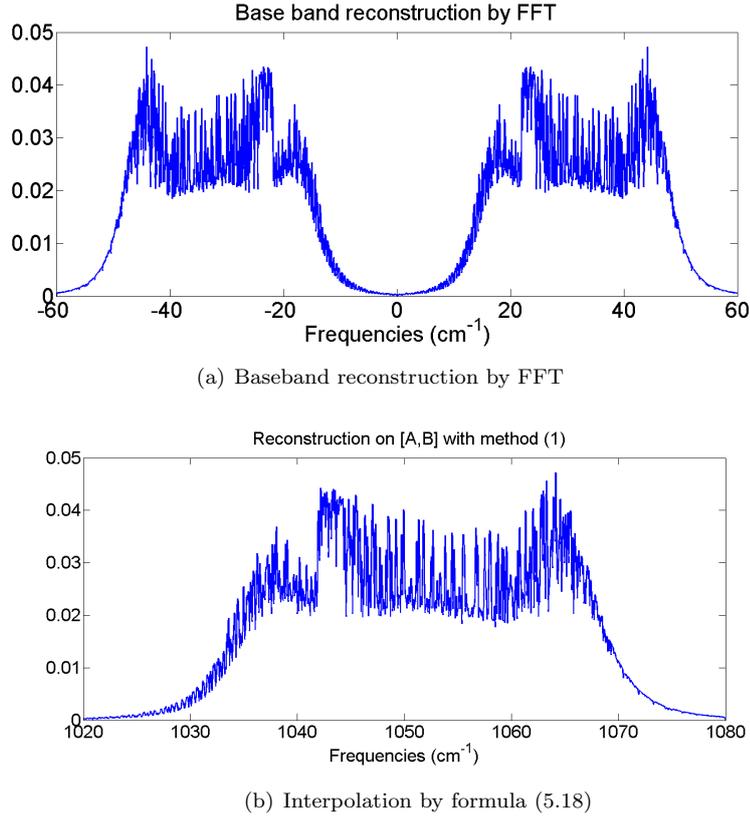


Figure 5.8: Reconstructions from the result of CNES method 1 (eq. (5.17)). The top reconstruction approximates a periodized version of the spectrum, whereas the bottom one is the more classical approximation on the interval  $[A, B]$ .

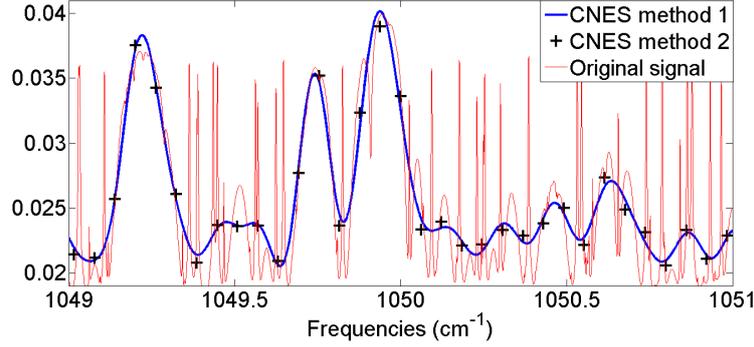


Figure 5.9: Results from the two state of the art methods (zoom around the  $1050 \text{ cm}^{-1}$  frequency). Because of the low  $OPD_{max} = 8 \text{ cm}$  compared to the spectrum  $\tilde{S}$  frequency contents, the reconstructions look like a low-pass version of the spectrum.

These two methods give an approximation of a certain vector, but the error cannot be made explicit easily as a function of  $S$  and the sampling parameters. We show in section 5.3.3 that the two methods are an approximation of a more general formulation using linear algebra and families of complex exponentials.

Our experiments are quite different from [94] since we are looking for a continuous approximation of the bandpass signal. Given interferometry measurements  $I_k = F(x_k)$ ,  $k = 1..M$ , we want to reconstruct the Fourier transform of  $F$  with support in  $\Omega$ . Approximation of  $F$  in a finite dimensional vectorial space can lead to acceptable results, as was done in CNES method 1 (5.17) with the family of functions  $\{\Phi(\cdot - l dx)\}_{l=0..N}$ . But assuming that  $F$  is a trigonometric polynomial as in [94] only gives a weak sense to  $\hat{F} = \tilde{S}$  as a sum of diracs. This happens somehow in CNES method 2 (5.19) where the only values of  $\tilde{S}_{rec}$  we know after reconstruction are the values at points  $\{A + l\Delta\sigma\}_{0 \leq l < \lfloor \frac{B-A}{\Delta\sigma} \rfloor}$ . This prevents any zoom or re-interpolation of the spectrum after reconstruction. To our view, this is incompatible with the type of information we have about  $\tilde{S}$  : scalar products with spatially badly localized function.

### 5.3.2 The general setting : nonharmonic Fourier series

We consider here an infinite number of samples : let  $\{y_k\}_{k \in \mathbb{Z}} \subset \mathbb{R}$  be a stable sampling set of  $BP(A, B)$ . The frame property (theorem 18) implies that the Fourier transform of  $f \in BP(A, B)$  is an infinite sum :

$$\hat{f}(\sigma) = \sum_{k \in \mathbb{Z}} c_k e^{iy_k \sigma} \quad \forall \sigma \in \Omega \quad (5.20)$$

with  $(c_k) \in l^2(\mathbb{Z})$ . The function  $f$  is obtained by pseudo-inversion of the sampling operator  $\mathcal{S}_\Lambda : f \rightarrow (f(y_k))_{k \in \mathbb{Z}}$  and the  $c_k$  coefficients are unique if and only if  $\{e^{i2\pi y_k \cdot}\}_{k \in \mathbb{Z}}$  is a Riesz basis of  $L^2(\Omega)$  (see [95] p. 169).

A non-stable sampling set which allows reconstruction is not valid here because the reconstruction operator would be unbounded (and so would be the reconstruction noise). Theorems 15, 16 and 17 give sufficient conditions on the set  $\Lambda$  to be a stable sampling set of  $BP(A, B)$ , this is why we will consider frames of complex exponentials of  $L^2(\Omega)$  of the form  $\{e^{i2\pi y_k \cdot}\}$  and approximate the unknown signal  $\tilde{S}$  as an infinite sum of these elements. Because  $F$  is symmetric we can assume that the  $y_k$  are nonnegative and replace complex exponentials by cosines :

$$\exists (c_k)_{k \in \mathbb{Z}} \in l^2(\mathbb{Z}) \quad \text{s.t.} \quad \tilde{S}(\cdot) = \sum_{k \in \mathbb{Z}} c_k \cos(2\pi y_k \cdot) \quad (5.21)$$

The data  $(I_k)_{k=1..M}$  rewrite

$$\begin{aligned} I_k &= \langle \tilde{S}(\cdot), e^{-i2\pi x_k \cdot} \rangle_{L^2(\Omega)} \\ &= \langle \tilde{S}(\cdot), \cos(2\pi x_k \cdot) \rangle_{L^2(\Omega)} \end{aligned}$$

since  $\tilde{S}$  is a real even function. Any approximation of  $\tilde{S}$  with a finite number of cosines leads to the resolution of the linear system

$$(I_k) = Q(c_l) \quad k = 1..M, l = 1..N \tag{5.22}$$

with  $Q_{k,l} = \langle \cos(2\pi y_k \cdot), \cos(2\pi x_k \cdot) \rangle_{L^2(\Omega)}$  and  $(c_l)$  the coefficients of the approximation in the basis

$\mathcal{B}_N = \{\cos(2\pi y_k \cdot)\}_{k=1..N}$ . The number of basis functions  $N$  can be chosen such that the system is invertible ( $N = M$ ), but the redundancy of the sampling at the neighbourhood of zero has an influence on the condition number. Again the resolution of the system with a thresholded pseudo-inversion is necessary.

### 5.3.2.1 Reconstruction is a projection operator

The following theorem states that for a given reconstruction frame  $\{\cos(2\pi y_k \cdot)\}_{k \in \mathbb{Z}}$ , and random sampling location  $\{x_k\}$ , the system (5.22) almost surely has a solution whenever  $N \geq M$ . When  $N = M$ , the reconstructed function is then the oblique projection of  $\tilde{S}$  on

$$Y_N = \langle \{\cos(2\pi y_l \cdot)\}_{l=1..N} \rangle$$

along the orthogonal of  $X_M = \langle \{\cos(2\pi x_k \cdot)\}_{k=1..M} \rangle$ .

**Theorem 20** *Existence of a solution Let  $\{y_l\}_{l \in \mathbb{Z}}$  be a subset of  $\mathbb{R}_+$  and  $\{x_k\}_{k=1..M}$  be independent real-valued random variables with absolutely continuous distribution for the Lebesgue measure. For all positive integer  $N \geq M$ , for all  $I \in \mathbb{R}^M$  the  $N \times M$  system*

$$I = Qc \tag{5.23}$$

$$\text{with} \quad Q_{k,l} = \langle \cos(2\pi x_k \cdot), \cos(2\pi y_l \cdot) \rangle_{L^2(\Omega)}$$

has almost surely at least one solution.

**Proof** We note  $Q(X)$  the  $N \times M$  matrix defined on  $\mathbb{R}^N$  by

$$(Q(X))_{k,l} = \langle \cos(2\pi X_k \cdot), \cos(2\pi y_l \cdot) \rangle_{L^2(\Omega)}$$

Given a sequence of strictly increasing positive numbers  $\{y_l\}_{l \in \mathbb{Z}}$ , the family  $\{\cos(2\pi y_l \cdot)\}_{l \in \mathbb{Z}}$  is a linearly independent family of functions on any interval  $[A, B]$  with  $A < B$  and thus on  $\Omega$ . Assume  $N \geq M$ , we take the first  $M$  functions from this family and prove that the restriction of  $Q(X)$  to its first  $M$  rows is almost surely invertible. We note  $Q_M$  this restriction with general term

$$\begin{aligned} (Q_M(X))_{k,l} &= \langle \cos(2\pi X_k \cdot), \cos(2\pi y_l \cdot) \rangle_{L^2(\Omega)} \\ &= \frac{1}{2}(\Phi(X_k + y_l) + \Phi(X_k - y_l)) \\ &= \Psi_l(X_k) \end{aligned}$$

$$\Phi = \mathcal{F}^{-1}(1_{\Omega}(\cdot))$$

We use the two lemmas below to prove that the zero set of the determinant of  $Q_M$  has measure zero.

**Lemma 1** Let  $\{\Psi_l\}_{l=1..M}$  be functions on  $\mathbb{R}$  and  $Q_M$  the  $M \times M$  matrix defined on  $\mathbb{R}^M$  by

$$(Q_M(X))_{k,l} = \Psi_l(X_k)$$

The determinant of  $Q_M$  is identically zero on  $\mathbb{R}^M$  if and only if the  $\{\Psi_l\}$  are linearly dependant on  $\mathbb{R}$

The proof of this lemma can be found in appendix 5.6.2. This implies that the determinant of  $Q_M$  is not identically zero on  $\mathbb{R}^M$ . Furthermore, it is a bandlimited function with spectral support in  $[-B, B]^M$  as a linear combination of tensor products of bandlimited functions. Lemma 2 implies that its zero level-set has zero measure in  $\mathbb{R}^M$ .

**Lemma 2** Measure of the zero set

Let  $f \in L^2(\mathbb{R}^M)$  with Fourier transform supported in a compact set.  $f$  is non identically zero if and only if its zero level set has zero Lebesgue measure on  $\mathbb{R}^M$ .

The proof is a classical result in Harmonic Analysis (see appendix 5.6.3). Finally  $Q_M(X)$  is invertible a.e. and the linear system (5.23) has almost surely a solution.

When  $M = N$ , the reconstruction  $\tilde{S}_{rec}$  is the projection of  $\tilde{S}$  on  $Y_N$  along  $(X_M)^\perp$  since it verifies for all  $0 \leq k \leq M$

$$\langle \tilde{S}_{rec}(\cdot), \cos(2\pi x_k \cdot) \rangle_{L^2(\Omega)} = \langle \tilde{S}(\cdot), \cos(2\pi x_k \cdot) \rangle_{L^2(\Omega)} \quad (5.24)$$

This oblique projection operator is not the best linear operator we can find which verifies this equation. The best reconstruction is the orthogonal projection of  $\tilde{S}$  on  $X_M$ .

**Theorem 21** Best reconstruction operator

Let  $P_M$  be the linear operator from  $L^2(\Omega)$  to  $\mathbb{C}^M$  defined by

$$P_M(S) = (\langle S, \cos(2\pi x_k \cdot) \rangle_{L^2(\Omega)})_{1 \leq k \leq M} \quad \forall S \in L^2(\Omega)$$

and  $\mathcal{K}$  the space of linear operators from  $\mathbb{C}^M$  to  $L^2(\Omega)$  such that

$$L \in \mathcal{K} \Leftrightarrow \langle LP_M(S), \cos(2\pi x_k \cdot) \rangle = I_k \quad k = 1..M$$

With the previous notations, the best reconstruction operator in terms of operator norm on  $\mathcal{K}$  is unique and gives the orthogonal projection on  $X_M$ .

$$L = \operatorname{argmin}_{L \in \mathcal{K}} \|Id - LP\| \quad \Leftrightarrow \quad LP = \operatorname{proj}_{\perp, X_M}$$

where the norm is the classical operator norm on the Hilbert space  $L^2(\Omega)$ .

The proof of this very intuitive result can be found in appendix 5.6.4. If we take the orthogonal projection as a reference, the reconstruction error  $\mathcal{E}$  is bounded by the difference between the oblique projection and the orthogonal projection on  $X_N$ , whose norm depends on the maximum angle  $\theta$  between a vector of  $X_M$  and a vector of  $Y_N$  (figure 5.10).

$$\begin{aligned} \|\mathcal{E}(\tilde{S})\| &\leq \|(\operatorname{proj}_{\perp, X_M} - \operatorname{proj}_{// (X_M)^\perp, Y_N})(\tilde{S})\| \\ &\leq \tan(\theta) \|\operatorname{proj}_{\perp, X_M}(\tilde{S})\| \\ &\leq \tan(\theta) \|\tilde{S}\| \end{aligned}$$

where  $\text{proj}_{\perp, X_M}$  holds for the orthogonal projection on  $X_M$  and  $\text{proj}_{// (X_M)^\perp, Y_N}$  is the oblique projection on  $Y_N$  along the  $X_M^\perp$  direction (this applies only in the case  $N = M$ ).

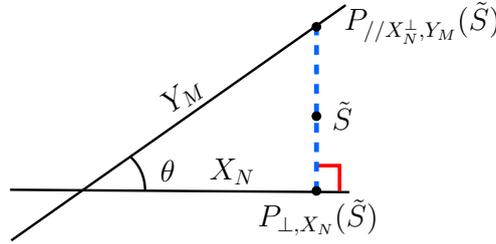


Figure 5.10: Oblique and orthogonal projections correspond to the two kind of approximations we get. The theoretical error is the distance between these two vectors, not the distance to  $\tilde{S}$ .

### 5.3.2.2 The bad condition number of the orthogonal projector

The computation of  $\text{proj}_{\perp, X_M}(\tilde{S})$  from the  $\{I_k\}$  is generally an ill-conditioned problem. When  $\{x_k\}_{k \in \mathbb{Z}}$  is a small perturbation of a tight frame, the proof of theorem 19 (appendix 5.6.1) shows that the reconstruction operator has a good behaviour, with a condition number smaller than  $(2e^{-2\pi BL} - 1)^{-1}$ , where  $L$  is the maximum perturbation (although the result holds in infinite dimension).

When the perturbation is larger than  $\frac{\ln(2)}{2\pi B}$ , or when the family of vectors  $\{\cos(2\pi x_k \cdot)\}_{k \in \mathbb{Z}}$  is not complete in  $L^2(\Omega)$ , the condition number can be very large. Figures 5.11(a) and 5.11(b) show the singular values and the deterioration in the condition number of the matrix  $Q$  with  $dx = dx_{crit} = 1/120$  and perturbations of standard deviation  $\nu$  in  $\{0, 1.10^{-5}, \dots, 5.10^{-4}\}$ . We also display the singular values in the case  $dx = 50 \mu m$  (figure 5.12) which verifies the conditions of theorem 15 and in the case of  $dx = 81 \mu m$  which does not verify this condition (figure 5.13). The first figure shows that the average number of invertible singular values is independant of the the sampling step  $dx$  and is equal to  $x_{max} \times \mu(\Omega) + 1 = 961$ . The second case illustrates the observation of Martin et al. [94] : when  $dx$  is forbidden by the generalized Shannon's theorem, perturbations increase the number of invertible singular values and thus the amount of information retrieved by the sampling/reconstruction couple.

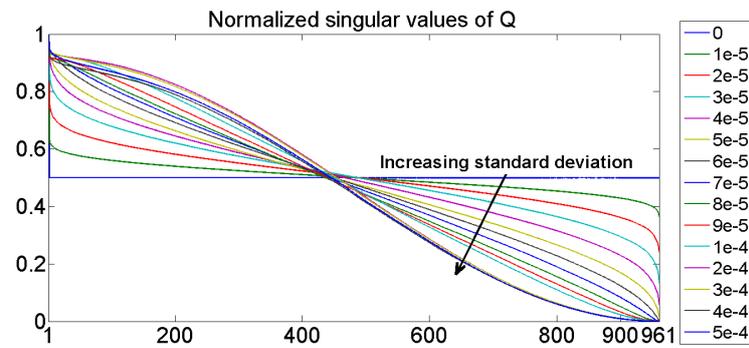
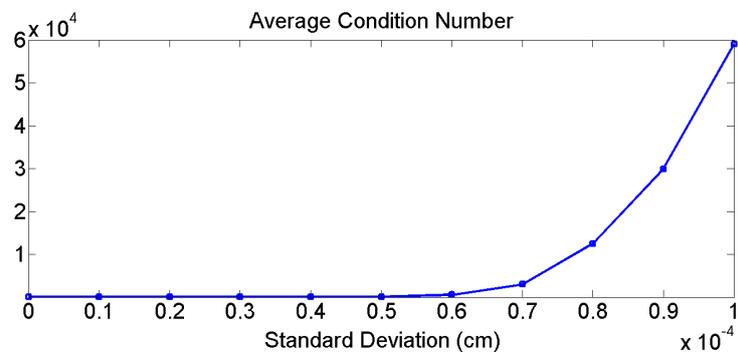
(a) Singular values of  $Q$ (b) Average condition number of  $Q$ 

Figure 5.11: Perturbed IES sampling in the orthogonal projection case. The standard deviation ranges from 0 to  $5 \cdot 10^{-4}$  cm. The larger the perturbation, the worse the condition number (average on 100 random simulations,  $dx = 83,3 \mu m$ ,  $OPD_{max} = 8$  cm).

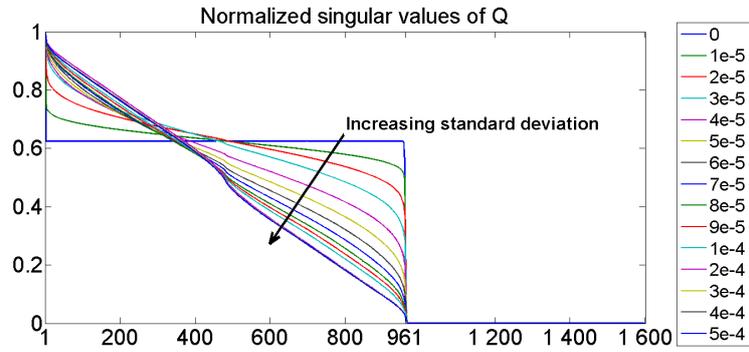


Figure 5.12: Perturbed IES sampling and orthogonal projection with  $dx = 50 \mu m$  : orthogonal projection. (average on 100 random simulations,  $OPD_{max} = 8 cm$ ). The perturbation in the sampling set deteriorates the conditioning of matrix  $Q$ .

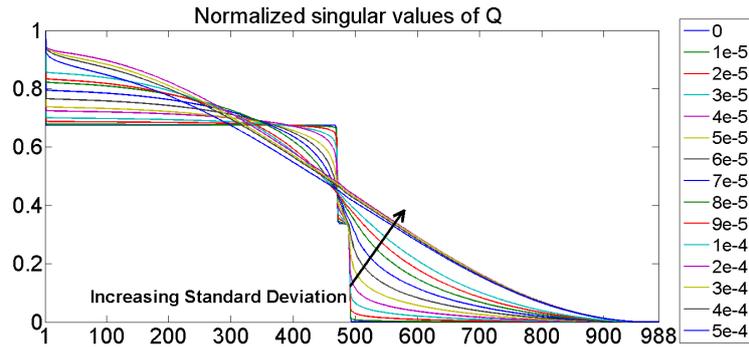
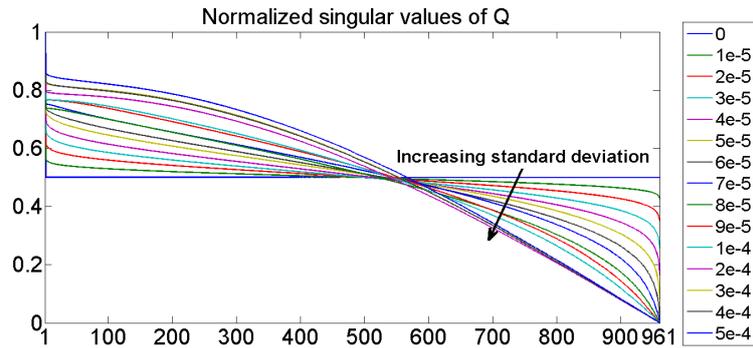


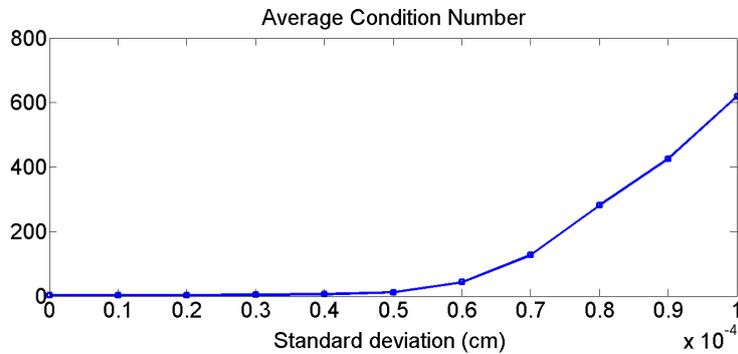
Figure 5.13: Perturbed IES sampling and orthogonal projection with  $dx = 81 \mu m$  (forbidden  $dx$ ) in the orthogonal projection case. (average on 100 random simulations,  $OPD_{max} = 8 cm$ ). The number of invertible singular values increases to approximately 965 as if the generalized Shannon's theorem were verified. Singular values over index 965 have a much lower amplitude.

### 5.3.2.3 Oblique projection

Any other choice of reconstruction basis  $\mathcal{B}_N$  leads to the oblique reconstruction operator if the number of reconstruction vectors  $N$  is equal to the number of samples  $M$ . This oblique projector can have a better condition number than the orthogonal projector. In figure 5.14(a) we display the singular values of operator  $Q$  with regular cosines reconstruction basis and the average condition number. This last one is smaller than in the irregular cosines case, with an average value of 600 with  $\nu = 1.0e - 4$  when the irregular cosines method is ten times worse. This explains the stability of this method and the differences in reconstructions from noisy data.



(a) Singular values of  $Q$



(b) Average condition number

Figure 5.14: Perturbed IES sampling with  $dx = 83.3 \mu m$  : oblique projection (average on 100 random simulations,  $OPD_{max} = 8 cm$ ). The reconstruction basis is  $\mathcal{B}_N = \{\cos(2\pi k dx \cdot)\}_{k=1..N}$ . Here the perturbation also deteriorates the conditioning, but the deterioration is less important.

### 5.3.3 A generalization of existing methods

We show here that the two cited methods (5.17) and (5.19) correspond to an approximation of the nonharmonic Fourier series method with certain basis functions of the form  $\cos(2\pi y_l \cdot)$ . The first

method is equivalent to the following assumption

$$\tilde{S}_{rec}(\cdot) = F(0) + 2 \sum_{k=1}^N F(l dx) \cos(2\pi l dx \cdot) \quad \text{on } \Omega$$

where the coefficients  $(F(l dx))_{l=1..N+1}$  are supposed to solve the linear system

$$(I_k) = Q F$$

with  $F_0 = F(0)$  and  $F_l = 2F(l dx)$  if  $l > 0$ . This is clearly the case  $y_l = l dx$  which we will call the *regular cosines* method.

The second method is only asymptotically similar to a nonharmonic Fourier series when the step  $\Delta\sigma$  tends to zero. Consider the matrix  $Q''$  as in (5.19) and note  $s_l = \tilde{S}_{rec}(A + l\Delta\sigma)$ . If  $\Delta\sigma$  is sufficiently small and if  $Q''$  has rank  $M$ , the pseudo-inverse of  $Q''$  has the expression

$$\begin{aligned} (Q'')^\dagger &= {}^t Q'' (Q'' {}^t Q'')^{-1} \\ &= \frac{1}{\Delta\sigma} {}^t Q'' \left( \frac{1}{\Delta\sigma} Q'' {}^t Q'' \right)^{-1}. \end{aligned}$$

The general term of  $\frac{1}{\Delta\sigma} Q'' {}^t Q''$  is a Riemann integral which tends to

$$\int_A^B \cos(2\pi x_k \sigma) \cos(2\pi x_l \sigma) d\sigma$$

when  $\Delta\sigma$  tends to zero and the general term of  $\frac{1}{\Delta\sigma} {}^t Q''$  is  $\cos(2\pi x_k (A + l\Delta\sigma))$ . Finally if  $Q$  is invertible and  $\Delta\sigma$  sufficiently small

$$s_l \approx \sum_{k=1}^M \cos(2\pi x_k (A + l\Delta\sigma)) Q^{-1} I.$$

The reconstructed vector  $s$  converges to the discretization on  $\{A + l\Delta\sigma\}$  of the solution of the nonharmonic Fourier method with  $y_l = x_l$ ,  $l = 1..M$ , which is the orthogonal projection of  $\tilde{S}$  on  $X_M$ . We call this method the *irregular cosines* method : the reconstruction basis and the measurement basis are the same.

### 5.3.4 Focus on Interlaced Sampling

#### 5.3.4.1 Non perturbed case

Interlaced sampling, also called periodic nonuniform sampling in the literature, combines two aliased signals in the case of two bands  $([-B, -A]$  and  $[A, B])$ . The more frequency bands the signal contains, the more shifted sampling grids are required for its reconstruction (see [62] [32] [52]). The following formulation only deals with two bands but can be extended to the  $n$ -bands case, even if they have different lengths.

Let  $f \in BP(A, B)$ ,  $0 < dx < \frac{1}{B-A}$ ,  $d > 0$ ,  $f_1, f_2 \in l^2(\mathbb{Z})$  defined by

$$f_1 = (f(k dx))_{k \in \mathbb{Z}} \quad , \quad f_2 = (f(k dx + d))_{k \in \mathbb{Z}}$$

By Poisson's Formula, the Fourier transforms, respectively  $F_1$  and  $F_2$ , verify

$$\begin{aligned} F_1(\cdot) &= \frac{1}{dx} \sum_{l \in \mathbb{Z}} \hat{f}(\cdot + \frac{l}{dx}) \\ F_2(\cdot) &= \frac{1}{dx} \sum_{l \in \mathbb{Z}} \hat{f}(\cdot + \frac{l}{dx}) e^{i2\pi d(\cdot + \frac{l}{dx})} \end{aligned} \tag{5.25}$$

Intervals  $[-B, -A]$  and  $[A, B]$  may overlap if  $dx$  does not verify the generalized Shannon's theorem. Consider an index  $k$  such that  $\Xi_k = ([-B, -A] + \frac{k}{dx}) \cap [A, B]$  has positive measure. We get the following system on  $\Xi_k$

$$\begin{cases} dx F_1(\cdot) &= \hat{f}(\cdot) + \hat{f}(\cdot - \frac{k}{dx}) \\ dx e^{-i2\pi d \cdot} F_2(\cdot) &= \hat{f}(\cdot) - e^{-i2\pi d \frac{k}{dx}} \hat{f}(\cdot - \frac{k}{dx}) \end{cases} \quad (5.26)$$

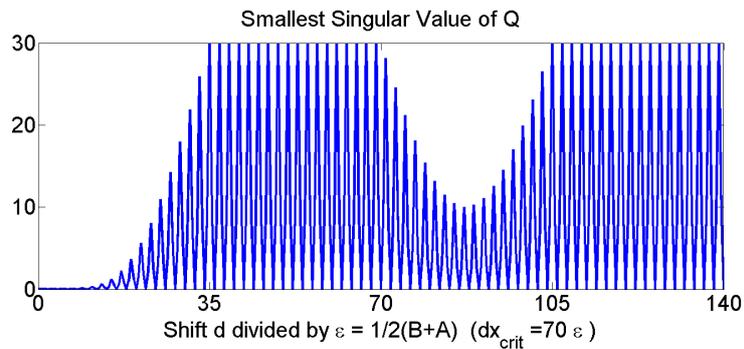
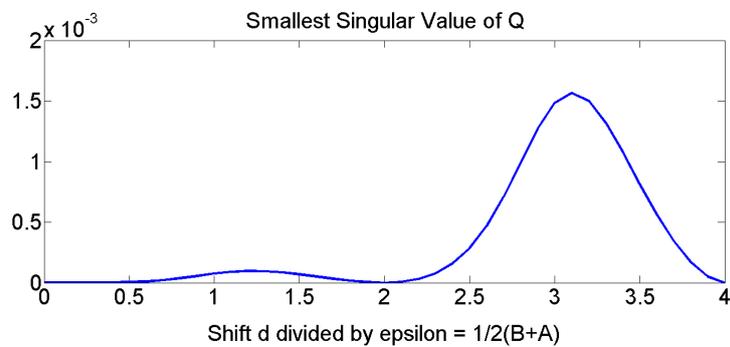
Since  $(a, b) \rightarrow (a, e^{i(2\pi d \xi)} b)$  is an isometry of  $\mathbb{C}^2$  we get the same condition as in (5.13) and we derive an energy estimate for the reconstructed signal on  $\Xi_k$  :

$$C^2 dx^2 (|F_1(\cdot)|^2 + |F_2(\cdot)|^2) \leq |\hat{f}(\cdot)|^2 + |\hat{f}(\cdot - \frac{k}{dx})|^2 \leq D^2 dx^2 (|F_1(\cdot)|^2 + |F_2(\cdot)|^2) \quad (5.27)$$

with  $C = \frac{1}{\sqrt{2+2|\cos(\pi d \frac{k}{dx})|}}$  and  $D = \frac{1}{2|\sin(\pi d \frac{k}{dx})|} \sqrt{2 + |2\cos(\pi d \frac{k}{dx})|}$  .

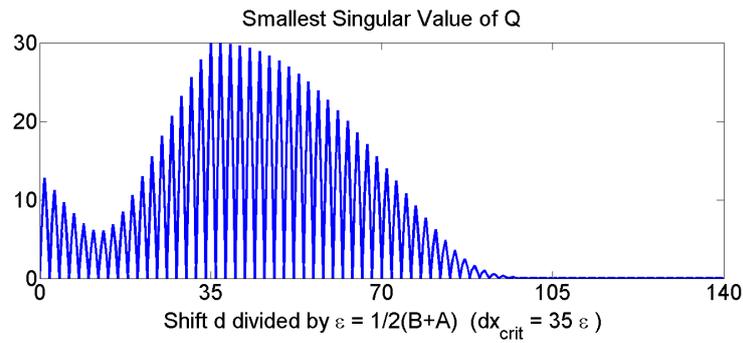
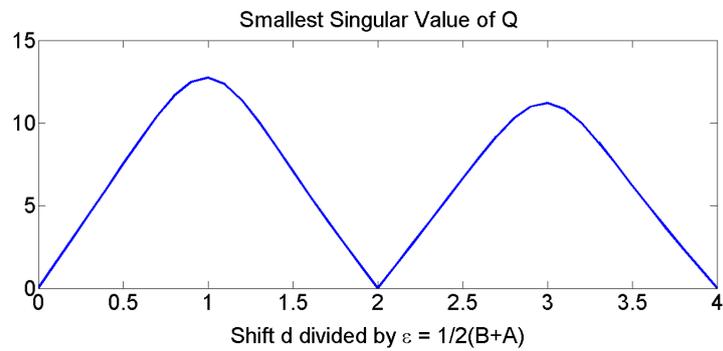
If  $0 < dx \leq \frac{1}{B-A}$  and  $(B+A) dx \in \mathbb{N}$  then there exists only one index  $k = (B+A) dx$  such that  $\mu(\Xi_k) > 0$  and the optimal shift corresponds to the values of  $d$  that minimize  $D$ , that is  $d \in \frac{dx}{2k} + \frac{dx}{k}\mathbb{Z} \Leftrightarrow d \in \frac{1}{2(B+A)} + \frac{1}{B+A}\mathbb{Z}$  (quadrature sampling).

This is true for an infinite sampling set but not necessarily for a finite number of samples, this is why we display in figures 5.15(a) and 5.16(a) the smallest singular value of  $Q$  as a function of the shift  $d$ . The two figures correspond to orthogonal projection (irregular cosines method) and oblique projection with the basis  $\{\cos(2\pi k dx_{crit} \cdot)\}$  and show that there indeed exist optimal values, close to  $\frac{1}{2(B+A)} + \frac{1}{B+A}\mathbb{Z}$ , and that the values in  $\frac{1}{B+A}\mathbb{Z}$  should not be used. These figures also show that the irregular cosines method is worse than the classical regular cosines method when  $d = 1/2(B+A)$ , that is quadrature sampling. The ratio is approximately  $1.7 \times 10^5$ , clearly the irregular cosines method cannot be used here with small shift values.

(a) Minimum singular value of  $Q$ .

(b) Zoom at origin of figure (a).

Figure 5.15: Non-perturbed case : smallest singular value of  $Q$  as a function of  $d$  (shift parameter) with irregular cosines method (orthogonal projection) and  $dx = 166.6 \mu m = 2 dx_{crit}$ ,  $OPD_{max} = 8 cm$ . It can take very small values compared to some other oblique projections (figure 5.16(b)).

(a) Minimum singular value of  $Q$ .

(b) Zoom at origin of figure (a).

Figure 5.16: Non-perturbed case : smallest singular value of  $Q$  as a function of  $d$  (shift parameter) with regular cosines method with  $dx = 166.6 \mu\text{m} = 2 dx_{crit}$ ,  $OPD_{max} = 8 \text{ cm}$ . Here the first extremum has amplitude  $\approx 13$  and the maximum value is 30 so the matrix  $Q$  is well-conditioned.

5.3.4.2 Perturbed case

In SIFTI interlaced sampling, the mirrors are covered with a deposit which adds a small length to the optical paths, dividing each facet into four smaller facets corresponding to four successive OPDs of the form  $\{x_k, x_k + \epsilon, x_k + 2\epsilon, x_k + 3\epsilon\}$ . We do not address the problem of a combination of four sampling grids and consider the approximation  $\{x_k\}_{k=1..N} \cup \{x_k + \epsilon\}_{k=1..N}$  (see section 5.1.1). Each position  $x_k$  is a perturbation of  $k dx$  where  $0 < dx \leq \frac{1}{B-A}$ . This sampling has less irregularity than the perturbed IES sampling described in the previous section. Of course the positions  $x_k$  and  $x_k + \epsilon$  are different from  $k dx$  and  $k dx + \epsilon$ , but the irregularity is the same in the two samples. Frame theory can be applied here to compute a bound on the perturbation as in theorem 19. Note that the existence of a positive bound  $L$  for the perturbation of a frame in bandlimited spaces is a very well known by Duffin and Shaeffer [41].

**Theorem 22** *Perturbed IIS Sampling*

Let  $0 < A < B$ ,  $0 < dx < \frac{\pi}{B-A}$  and  $d$  satisfying theorem 16. There exists a constant  $L > 0$  such that the set  $\Lambda = \{x_k\}_{k \in \mathbb{Z}}$  which verifies

$$\begin{cases} |x_{2k} - k dx| \leq L \\ |x_{2k+1} - k dx - d| \leq L \end{cases} \quad \forall k \in \mathbb{Z}$$

is a stable sampling set of  $BP(A, B)$ . When  $dx > \frac{1}{A+B}$  and  $(B + A) dx \in \mathbb{N}$  (exact overlapping in Fourier domain), the constant  $L$  is at least  $\frac{\ln(1+1/\kappa)}{2\pi B}$ , where  $\kappa$  is the condition number of the  $2 \times 2$  matrix in equation (5.13)

$$\kappa = \text{cond}\left(\begin{pmatrix} 1 & 1 \\ 1 & e^{-i2\pi d \frac{k}{dx}} \end{pmatrix}\right)$$

The computation of this bound can be found in appendix 5.6.5. One interesting remark is that this bound can never be larger than  $d/2$ , which prevents samples superposition and loss of information : it ensures that the set  $\Lambda$  is uniformly discrete.

This bound is certainly not optimal when  $d > \frac{1}{2(B+A)}$  because of the periodicity of the condition number  $\kappa(d)$  at fixed  $dx$ . For  $d = \frac{1}{2(A+B)}$  the bound  $L$  is  $\frac{\ln(2)}{2\pi B}$ , as in theorem 19.

As shown in figure 5.16(a) the IIS sampling together with regular cosines has a good numerical behaviour for some values in  $\frac{1}{4f_c} + \frac{1}{2f_c}\mathbb{Z}$  but not better than non-perturbed IES sampling (figure 5.14(a) ). The main advantage of IIS sampling is its stability to perturbations in the OPDs  $x_k$  as shown in figure 5.17. On the contrary of perturbed IES sampling (figure 5.14(a)) the smallest singular value at  $d \notin \frac{1}{2f_c}\mathbb{Z}$  does not tend to zero as the standard deviation increases.

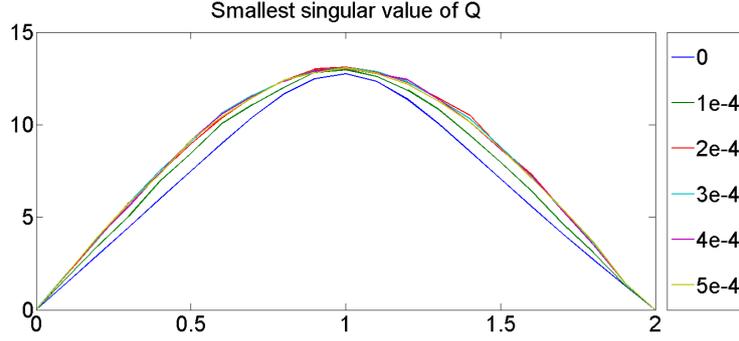


Figure 5.17: IIS with random perturbations : regular cosines method (the standard deviation is displayed on the right). The reconstruction basis is  $\mathcal{B} = \{\cos(2\pi k dx_{crit} \cdot)\}$  and  $OPD_{max} = 8 \text{ cm}$ . The quadrature sampling (extremum) also corresponds to a well-conditioned matrix.

We did not find any reference to this stability of interlaced sampling in the literature. This is a very remarkable property of multiband signals and multichannel sampling.

## 5.4 Link with the dual basis

Choosing the same reconstruction basis and sampling basis guaranties the best operator error norm among all possible linear operators but does not guaranty the good numerical behaviour. In the previous paragraphs the orthogonal projection had an acceptable condition number in only one case : the regular uniform sampling. This is due to the fact that the regular cosines basis (for  $dx = dx_{crit}$ ) is orthonormal, and thus its own dual. The dual basis of  $\mathcal{B} = \{g_k\}_{k \in \mathbb{Z}}$  is the set of functions  $\tilde{\mathcal{B}} = \{\tilde{g}_k\}_{k \in \mathbb{Z}}$  such that

$$\langle g_k, \tilde{g}_l \rangle = \delta_{k,l} \quad \forall k, l \in \mathbb{Z} \quad (5.28)$$

and any signal can be reconstructed with the formula

$$f(\cdot) = \sum_{k \in \mathbb{Z}} \langle f, g_k \rangle \tilde{g}_k$$

Choosing the dual basis as the reconstruction basis is then the easiest way to solve our problem, since the matrix  $Q$  is now identity. In terms of inversion norm, the class of optimal reconstruction bases is the dual basis, up to an isometry. In the following we propose an explanation for the bad behaviour of the irregular cosines method in periodic nonuniform sampling. We consider the case  $(B + A)dx = k \in \mathbb{N}$  as in section 5.3.4, and rewrite the equations verified by  $\hat{f}(\cdot)$ ,  $\hat{f}(\cdot + \frac{k}{dx})$ ,  $F_1$  and  $F_2$  on  $[A, B]$

$$\begin{pmatrix} \hat{f}(\cdot) \\ \hat{f}(\cdot - \frac{k}{dx}) \end{pmatrix} = \alpha \begin{pmatrix} e^{-i2\pi d \frac{k}{dx}} & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} F_1(\cdot) \\ e^{-i2\pi d \cdot} F_2(\cdot) \end{pmatrix} \quad (5.29)$$

with  $\alpha(d) = dx / (e^{-i2\pi d k/dx} - 1)$ . Taking the inverse Fourier transform of  $\hat{f}(\cdot)$  and  $\hat{f}(\cdot - (B - A))$  on  $[A, B]$  gives a decomposition of  $f$  as in (5.14) from which we derive  $\widehat{\Phi}_1$  and  $\widehat{\Phi}_2$  on  $[A, B]$ .

$$\begin{aligned} \alpha(d)^{-1} \widehat{\Phi}_1(\cdot) &= \begin{cases} -1 & \text{on } [-B, -A] \\ e^{-i2\pi d k/dx} & \text{on } [A, B] \end{cases} \\ \alpha(d)^{-1} \widehat{\Phi}_2(\cdot) &= \begin{cases} 1 & \text{on } [-B, -A] \\ -1 & \text{on } [A, B] \end{cases} \end{aligned} \tag{5.30}$$

We are looking for an even spectrum  $\tilde{S}$  and thus consider the even parts  $\Phi_{1,l}^e$  and  $\Phi_{2,l}^e$  of the Fourier transforms of  $\Phi_1(\cdot - l dx)$  and  $\Phi_2(\cdot - l dx - d)$  :

$$\begin{cases} \alpha(d)^{-1} \widehat{\Phi}_{1,l}^e(\cdot) = -i e^{-i2\pi d f_c} \sin(2\pi(l dx \cdot + d f_c)) \\ \alpha(d)^{-1} \widehat{\Phi}_{2,l}^e(\cdot) = i e^{-i2\pi d f_c} \sin(2\pi((l dx + d) \cdot - d f_c)) \end{cases} \tag{5.31}$$

and now taking the real parts, since the spectrum we seek is real, we get the correct reconstruction basis in the Fourier domain

$$\begin{cases} \widehat{\Psi}_{1,l}(\cdot) = \frac{dx}{2 \sin(2\pi d f_c)} \sin(2\pi(l dx \cdot + d f_c)) \\ \widehat{\Psi}_{2,l}(\cdot) = \frac{dx}{2 \sin(2\pi d f_c)} \sin(2\pi(d f_c - (l dx + d) \cdot)) \end{cases} \tag{5.32}$$

We notice that  $d \in \frac{1}{4f_c} + \frac{1}{2f_c} \mathbb{Z}$  is a special case because the  $\widehat{\Psi}_{1,l}$  functions become  $\frac{dx}{2} \cos(2\pi l dx \cdot)$  and the  $\widehat{\Psi}_{2,l}$  become  $\frac{dx}{2} \cos(2\pi(l dx + d) \cdot)$ .

This is the irregular cosines basis from figure 5.15(a), but since we took the even part of the  $\Psi_{1,l}$  and  $\Psi_{2,l}$ , the duality relation (5.28) is no more verified. A closer look at the whole set of singular values of  $Q$  when  $d \in \frac{1}{4f_c} + \frac{1}{2f_c} \mathbb{Z}$  reveals that this is indeed a good approximation basis as nearly all the singular values are constant and equal to  $\frac{1}{dx}$ . We display in figure 5.18 the singular values in the perturbed case : the approximation basis is fixed and equal to  $\{\Psi_{1,l}\} \cup \{\Psi_{2,l}\}$  (up to a  $\frac{dx}{2}$  factor) with  $d = \frac{1}{4f_c}$ .

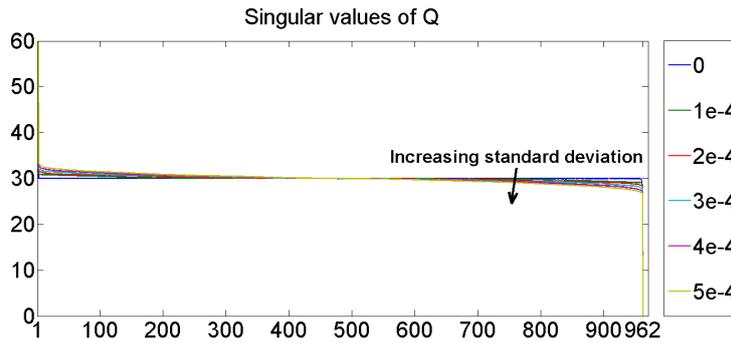


Figure 5.18: Singular values of  $Q$  in quadrature IIS sampling with perturbation (displayed on the right). Although the last singular value is small, all the others are approximately equal to  $\frac{1}{dx} = 30$  and very stable with the perturbation.

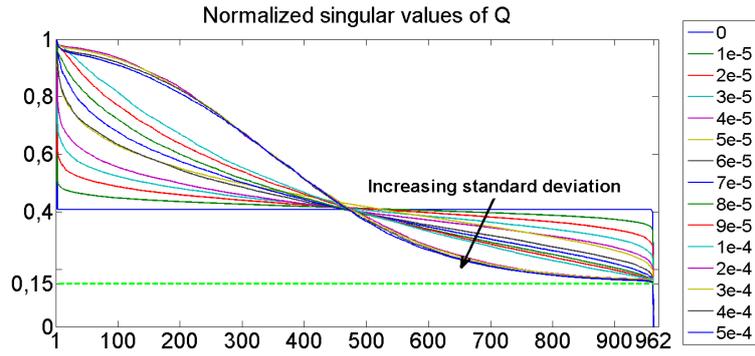
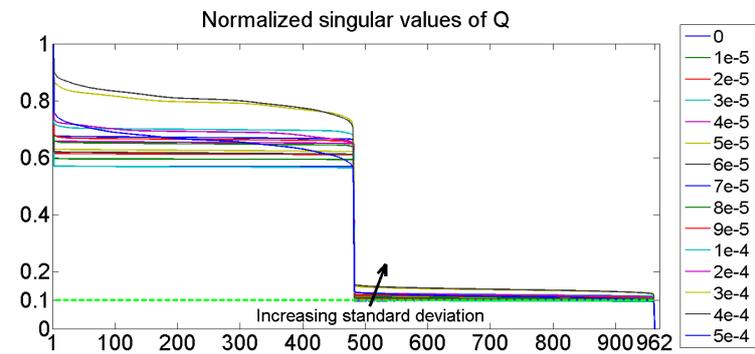
(a)  $\mathcal{B}_1 = \{\Psi_{1,l}\} \cup \{\Psi_{2,l}\}$  (fixed) : oblique projection.(b)  $\mathcal{B}_2 = \{\cos(2\pi x_k \cdot)\}$  : orthogonal projection.

Figure 5.19: Singular values of  $Q$  in IIS sampling with different approximation bases. Here the shift parameter is half the optimal parameter ( $d = \frac{1}{4f_c}$ ). The singular values are still bounded by below as the perturbation increases.

The behaviour in the non-quadrature case is similar but slightly different. We display in figures 5.19 the singular values of  $Q$  when  $d = \frac{1}{4f_c}$  along with the perturbation amplitude. In both cases all the singular values but the last one are bounded by below but the bound decreases in the first case (figure 5.19(a)) and slowly increases in the orthogonal projection case (figure 5.19(b)). Both methods are stable to noise in the interferogram thanks to this bound, so we can conclude that perturbed IIS sampling is clearly more stable than perturbed IES sampling for a large number of reconstruction bases.

## 5.5 Conclusion

In this chapter we addressed the problem of restoring a spectrogram (*i.e.* the Fourier transform of a real-symmetric narrow-band limited signal<sup>2</sup>) from its interferogram measured at slightly perturbed regular samples. We studied two methods previously proposed by CNES, in the framework of the SIFTI project. We further developed the theory allowing to correct numerical problems and understand their properties in terms of approximation and stability, depending on the sampling geometry.

Concerning sampling modes, we considered first the infinite dimensional case (infinite number of samples), which allows to determine the cases where perfect reconstruction is possible. Four situations were analysed: Periodic IES, *i.e.* regular sampling, Periodic IIS, *i.e.* the union of a regular sampling grid and its shift by a constant, and their (non-periodic) randomly perturbed counterparts. In the periodic sampling cases, IES critical sampling is governed by the generalized Shannon sampling theorem, and IIS by its nonuniform counterpart, both well known from early developments in telecommunications theory (see [57], [92] and [62]). The perturbed counterparts were much less studied, and were the subject of the section 5.3. First the Duffin-Schaeffer theorem for perturbed (IES) sampling of band-limited signals is extended to the case of narrow-band signals (theorem 19), then this result is extended to the perturbed IIS case (theorem 22). In both cases an upper bound  $\frac{\ln 2}{2\pi B}$  on the perturbation is established, which provides a sufficient condition for stable reconstruction. However, whether this bound is tight is still an open problem, since no counterexample could be found for a perturbation larger than this upper bound and smaller than the obvious upper bound provided by half the sampling step. Furthermore this upper bound is about five times smaller than the perturbations that arise in practice due to limited manufacturing precision within the SIFTI project.

In practice only a very limited number of samples is allowed, whereas the support of the interferogram is either infinite or at least much larger (up to noise precision) than the attainable range. Hence perfect reconstruction is not possible, and we can only expect to obtain good approximations which are stable to noisy measurements. In fact we show, under a very general setting, two theoretical results that correspond to these two optimization criteria. We consider the vector space  $X_M$  spanned by the measurement functions and the vector space  $Y_N$  spanned by the reconstruction functions. Then: *(i)* The best approximation (in the sense of min-max  $L^2$  error under interpolation constraints) corresponds to the orthogonal projection on  $X$  (see theorem 7); *(ii)* Given a set of reconstruction functions spanning  $Y_M$ , the most stable approximation (in the sense that the condition number of the reconstruction operator is minimized) for a given reconstruction space is given by the oblique projection on  $Y_M$  and orthogonal to  $X_N$ .

We observed that the second method proposed by CNES (pseudo-inverse) actually corresponds to a numerical approximation to case *(i)* where the matrix of dot-products between the measurement and reconstruction functions is approximated by Riemann sums. When substituting these Riemann sums by the exact analytical dot-products we obtain the desired best approximation reconstruction.

In addition the first method proposed by CNES actually corresponds to case *(ii)* with  $X_N$  the irregular cosines measurement functions, and  $Y_M$  the regular cosines reconstruction functions, which explains its relatively good conditioning.

A further increase in conditioning is still possible if we choose  $Y_M$  to be spanned by the dual basis of the measurement functions spanning  $X_N$ . However, as pointed out in section 4, the analytical computation of this basis is non trivial, and is left as an open problem for future research.

---

<sup>2</sup>*i.e.* with spectral support in  $[A, B] \subset [0, \infty[$

Future research could also be conducted in order to devise a method that makes a good compromise between best approximation and stability. In this context IIS sampling is a very interesting scheme as it was shown stable to noise in the data (singular values are bounded by below, except the last one) for several reconstruction bases.

## 5.6 Appendix

### 5.6.1 Perturbation constant $\frac{\ln(2)}{2\pi B}$

Let  $dx > 0$  verifying the generalized Shannon's condition (5.11). For all  $f \in BP(A, B)$  we have the Poisson's formula

$$\sum_{k \in \mathbb{Z}} f(k dx) e^{i2\pi k dx \cdot} = \frac{1}{dx} \sum_{k \in \mathbb{Z}} \hat{f}\left(\cdot + \frac{k}{dx}\right) \quad \text{a.e.}$$

and since there is no overlapping, the integral of the periodized Fourier transform over an interval of length  $\frac{1}{dx}$  is equal to the integral on  $\Omega$ . By Plancherel's formula we get

$$\begin{aligned} \|f(k dx)\|_{l^2(\mathbb{Z})}^2 &= \frac{1}{dx} \|\hat{f}\|_{L^2(\Omega)}^2 \\ &= \frac{1}{dx} \|f\|_{L^2(\Omega)}^2 \end{aligned}$$

$$\text{with} \quad f(k dx) = \left\langle \hat{f}(\cdot), e^{i2\pi k dx \cdot} \right\rangle_{L^2(\Omega)}$$

so  $\{e^{i2\pi k dx \cdot}\}_{k \in \mathbb{Z}}$  is a tight frame of  $L^2(\Omega)$ . Since  $BP(A, B)$  is stable under differentiation, we have for any integer  $l \geq 1$

$$\begin{aligned} \|f^{(l)}(k dx)\|_{l^2(\mathbb{Z})}^2 &= \frac{1}{dx} \|\widehat{f^{(l)}}\|^2 \\ &= \frac{1}{dx} \|(i2\pi \cdot)^l \hat{f}(\cdot)\|^2 \\ &\leq \frac{(2\pi B)^{2l}}{dx} \|f\|^2 \end{aligned}$$

We then use similar arguments as [41], and develop  $(f(x_k) - f(k dx))$  with Taylor series :

$$f(x_k) - f(k dx) = \sum_{l \geq 1} \frac{1}{l!} f^{(l)}(k dx) (x_k - k dx)^l$$

By Cauchy-Schwartz inequality (taking  $\alpha > 0$ ) :

$$\begin{aligned} &\sum_{k \in \mathbb{Z}} |f(x_k) - f(k dx)|^2 \\ &= \sum_{k \in \mathbb{Z}} \left| \sum_{l \geq 1} \frac{1}{l!} f^{(l)}(k dx) (x_k - k dx)^l \right|^2 \\ &\leq \sum_{k \in \mathbb{Z}} \left( \sum_{l \geq 1} \frac{\alpha^{2l}}{l!} |f^{(l)}(k dx)|^2 \right) \left( \sum_{l \geq 1} \frac{1}{l!} \left(\frac{L}{\alpha}\right)^{2l} \right) \\ &\leq \frac{1}{dx} \|f\|^2 \left( e^{\alpha^2 (2\pi B)^2} - 1 \right) \left( e^{\alpha^{-2} L^2} - 1 \right) \end{aligned}$$

which is minimal for  $\alpha = \sqrt{\frac{L}{2\pi B}}$  :

$$\Rightarrow \sum_{k \in \mathbb{Z}} |f(k dx) - f(x_k)|^2 \leq \frac{1}{dx} (e^{2\pi B L} - 1)^2 \|f\|^2$$

and by triangular inequality :

$$\frac{1}{\sqrt{dx}} (2 - e^{2\pi B L}) \|f\| \leq \sqrt{\sum |f(x_k)|^2} \leq \frac{1}{\sqrt{dx}} e^{2\pi B L} \|f\|$$

A sufficient condition for having a frame is thus  $L < \frac{\ln(2)}{2\pi B}$ .

### 5.6.2 Non zero determinant

We proceed by induction. In dimension 1 the linear dependence of  $\{\psi_1\}$  means that there exist  $a_1 \neq 0$  such that  $a_1\psi_1 = 0$  on  $\mathbb{R}$  which is equivalent to a zero determinant on  $\mathbb{R}$  of the  $1 \times 1$  matrix  $(\Psi_1(\cdot))$ . Assume that the proposition is true in dimension  $M \geq 1$ .

Let  $\{\Psi_l\}_{l=1..M+1}$  be linearly independent functions on  $\mathbb{R}$  and  $Q$  be the  $(M+1) \times (M+1)$  matrix with general term  $Q_{k,l}(x) = \Psi_l(x_k)$ . The determinant of  $Q$  can be developed with Laplace's formula :

$$\det(Q(x)) = \sum_{l=1}^{M+1} (-1)^{l+M+1} \Psi_l(x_{M+1}) \det((Q \setminus T_{n+1,l})(x))$$

where  $T_{n+1,l}$  is the union of the last row and  $l$ -th column and the sub-determinants do not depend on  $x_{M+1}$ . If  $\det(Q(x))$  is identically zero on  $\mathbb{R}^{M+1}$  then by linear independence of the  $\{\Psi_l\}$  the sub-determinants are identically zero which is in contradiction with the recurrence hypothesis. The converse is trivial (for all  $M \geq 1$ , if the  $\{\Psi_l\}_{l=1..M}$  are linearly dependant then the determinant  $\det(Q(x))$  is zero everywhere).

### 5.6.3 Measure of the zero set

This is a classical result with trigonometric polynomials [6] and the proof is similar with bandlimited functions. In dimension 1 the zeros of a bandlimited function are uniformly discrete by the analytic extension theorem and isolated zeros theorem. We proceed by induction : let  $\text{supp}(\hat{f}) \subset [-B, B]^{M+1}$  with  $M \geq 1$ . The function  $f(x_1, \dots, x_M, \cdot)$  is bandlimited for all  $(x_1, \dots, x_M) \in \mathbb{R}^M$  so its zero set has measure zero. The zero set of  $f$ , denoted by  $Z_f$ , has finite measure on any compact subset  $K$  of  $\mathbb{R}^{M+1}$ . The indicator function of  $Z_f$  is measurable and belongs to  $L^1_{loc}(\mathbb{R}^{M+1})$ . By Fubini's theorem we have

$$\begin{aligned} \mu(Z_f \cap K) &= \int_{\mathbb{R}^M} \int_{\mathbb{R}} \mathbb{1}_{Z_f \cap K}(x_1 \dots x_M, y) dy dx_1 \dots dx_M \\ &= \int_{\mathbb{R}^M} 0 dx_1 \dots dx_M \\ &= 0 \quad \forall K \subset \subset \mathbb{R}^{M+1} \end{aligned}$$

By upper continuity of the measure, and considering an exhaustive sequence of compacts on  $\mathbb{R}^{M+1}$ , we get  $\mu(Z_f) = 0$ .

### 5.6.4 Best reconstruction operator

We show here that the minimizer of  $\|Id - LP\|_{2,2}$  among all linear operators from  $\mathbb{R}^M$  to  $L^2(\Omega)$  such that

$$\langle LP S, \cos(2\pi x_k \cdot) \rangle = I_k \quad \forall S \in L^2(\Omega)$$

is unique and verifies  $LP = \text{proj}_{\perp, X_M}$ . The above equation can be reformulated :

$$P L P = P \quad \text{on } L^2(\Omega)$$

Since  $\text{Ker}(P) = X_M^\perp$ , we get

$$\begin{aligned} P(LP - \text{proj}_{\perp, X_M}) &= P(Id - \text{proj}_{\perp, X_M}) \\ &= P \text{proj}_{\perp, X_M^\perp} \\ &= 0 \end{aligned}$$

Thus,  $R(LP - proj_{\perp, X_M}) \subset X_M^{\perp}$  which we rewrite

$$LP = proj_{\perp, X_M} + TP$$

with  $R(T) \subset X_M^{\perp}$ . If  $T = 0$ ,  $LP$  is the projection on  $X_M$  and  $Id - LP = proj_{\perp, X_M^{\perp}}$  has norm 1. If  $T \neq 0$ , then we show that  $Id - LP$  has a norm  $> 1$ . Consider  $f \in X_M$  such that  $TP(f) \neq 0$  and  $\|f\| \in ]0, 1]$ . Then consider  $g = \lambda TP(f)$  with  $\lambda = -\frac{\sqrt{1-\|f\|^2}}{\|TP(f)\|}$  such that  $\|f\|^2 + \|g\|^2 = 1$ . We get

$$\begin{aligned} & (Id - LP)(f + g) \\ &= f + g - proj_{\perp, X_M}(f + g) - TP(f + g) \\ &= f + g - f - TP(f) \\ &= g - TP(f) \end{aligned}$$

$$\|(Id - LP)(f + g)\|^2 = (\|TP(f)\| + \sqrt{1 - \|f\|^2})^2$$

this is also true for  $\alpha f$  with  $0 < \alpha \leq 1$  and we develop this expression when  $\alpha \rightarrow 0$  :

$$\|(Id - LP)(f)\| \approx 1 + \alpha \|TP(f)\|$$

which implies  $\|Id - LP\| > 1$ . The minimum of  $\|Id - LP\|$  on  $\mathcal{K}$  is 1 and is obtained for  $L$  verifying  $LP = proj_{\perp, X_M}$ .

Uniqueness : assume that two operators  $L_1$  and  $L_2$  in  $\mathcal{K}$  verify this equation, by linearity :

$$(L_1 - L_2)P = 0 \quad \text{on } L^2(\Omega)$$

Since  $R(P) = \mathbb{C}^M$ ,  $L_1 - L_2$  is the null operator on  $\mathbb{C}^M$  and  $L_1 = L_2$ .

### 5.6.5 Perturbation constant $\frac{\ln(1+1/\kappa)}{2\pi B}$

The existence of the perturbation bound  $L > 0$  results from the frame property of the family of functions  $\mathcal{G} = \{e^{i2\pi k dx}\}_{k \in \mathbb{Z}} \cup \{e^{i2\pi(k dx + d)}\}_{k \in \mathbb{Z}}$  in  $L^2(\Omega)$  (see [41] Lemma II). In the special case when  $(B + A)dx = k > 1$  with  $k \in \mathbb{N}$

$$[-A, -B] + \frac{k}{dx} = [A, B]$$

we get the inequalities (5.27), and by integration on  $[A, A + \frac{1}{dx}]$

$$\begin{aligned} C^2 dx (\|F_1\|^2 + \|F_2\|^2) &\leq \|\hat{f}\|^2 \leq D^2 dx (\|F_1\|^2 + \|F_2\|^2) \\ \Leftrightarrow \frac{1}{D^2 dx} \|f\|^2 &\leq \sum_{k \in \mathbb{Z}} |f(x_k)|^2 \leq \frac{1}{C^2 dx} \|f\|^2 \end{aligned}$$

since  $\hat{f}$  is zero on  $[B, A + \frac{1}{dx}]$  and  $[-A, -B + \frac{1}{dx}]$ . These inequalities imply that the frame bounds of  $\mathcal{G}$  are  $\frac{1}{D^2 dx}$  and  $\frac{1}{C^2 dx}$ . Following the same scheme as in [41] we get  $L < \frac{\ln(1+C/D)}{2\pi B}$  where  $\frac{D}{C}$  is the condition number  $\kappa$ .

Remark : the set  $\Lambda$  is uniformly discrete (there exists  $e > 0$  such that  $|x_m - x_n| > e$  whenever  $m \neq n$ ). To show this, consider the derivative of  $L_{max}(d) = \frac{\ln(1+C(d)/D(d))}{2\pi B}$ . Then

$$L'_{max}(d) = \frac{f_c}{B} \frac{1}{(1 + \cos(2\pi f_c d) + \sin(2\pi f_c d))}$$

for  $d \in [0, \frac{1}{4f_c}]$ . Furthermore  $L_{max}(\cdot)$  is even and periodic with period  $\frac{1}{2f_c}$ , so the differences of  $L_{max}$  are bounded.

$$L_{max}(d) \leq \frac{f_c}{B} \frac{|d - \frac{k}{2f_c}|}{2} \quad \text{for } d \in \left[-\frac{1}{4f_c}, \frac{1}{4f_c}\right] + \frac{k}{2f_c}$$

$$\Rightarrow \quad x_{2k+1} - x_{2k} \geq d - \frac{f_c}{B} \left|d - \frac{k}{2f_c}\right|$$

and by a shift in indices we also have

$$x_{2k+2} - x_{2k+1} \geq dx - d - \frac{f_c}{B} \left|dx - d + \frac{k'}{2f_c}\right|$$

where  $dx - d \in \left[-\frac{1}{4f_c}, \frac{1}{4f_c}\right] + \frac{k'}{2f_c}$ .

## Chapter 6

# Conclusion

The large scope of this thesis, initially devoted to an image processing problem (chapter 4), reveals the importance of adapted mathematical tools in satellite remote sensing. The usual framework of Fourier's theory is often not sufficient, more general modellings and more evolved algorithms are necessary to efficiently handle the irregular sampling phenomenon. This latest raises both theoretical (well-posedness and stability of such samplings) and numerical questions (choice of an approximation space, regularization, boundary conditions, computational complexity) as we observed in the three major problems we treated.

Our contributions to the scientific research through this thesis first showed the feasibility and accuracy of fast irregular sampling restoration in image processing thanks to spline functions, non-quadratic regularization and dual approaches (chapter 4).

In chapter 3 we considered a highly ill-posed problem in stereoscopy and showed the good performances of modern  $\ell^1$  and  $\ell^0$  approaches with ISTA and IHTA algorithms while classical approaches completely fail. The model has been later completed with an assumption on the elevation which allows for more accurate corrections in the case of correlated elevation and microvibrations.

And last we studied the problem of spectrum reconstruction from irregular interferometry measurements and obtained theoretical results on the well-posedness of the problem and the allowed maximal perturbation. Our modelling is a generalization of existing methods and provides a better understanding of the sampling set properties in bandpass spaces.

A general principle in this work on applied problems involving irregular sampling sets is the importance of a good modelling and its influence on the implementation of sampling operators and algorithms. The behaviour of the restored signals is directly related to this modelling and this thesis has been an exciting journey to the land of sparse signals, stair-cases, splines, wavelets, Fourier atoms, convex sets, hyper-planes and many others. The more you know them and their capabilities, the more possibilities you are offered in the resolution of high level applied problems.

As a perspective, the simultaneous resolution of the microvibrations and image restoration problems is a natural continuation of this work. More complex models including some unconsidered aspects of the physical phenomena could be examined such as the spatially variant TDI blurring kernel or the irregular sampling in the disparity map and also the more accurate TV-TV2 inf-convolution model for disparity data. I hope this thesis work will encourage researchers to consider these promising improvements.



# Bibliography

- [1] A. Aldroubi, M. Unser, and M. Eden. Cardinal spline filters: Stability and convergence to the ideal sinc interpolator. *Signal Processing*, 28(2):127–138, August 1992.
- [2] A Almansa. Echantillonnage, interpolation et détection. applications en imagerie satellitaire., 2002.
- [3] A. Almansa, J. Caron, and S. Durand. Deblurring of irregularly sampled images by TV regularization in a spline space. *IEEE Proceedings of the 17<sup>th</sup> International Conference on Image Processing*, pages 1181 – 1184, 2010.
- [4] Andrés Almansa, Vicent Caselles, Gloria Haro, and Bernard Rougé. Restoration and zoom of irregularly sampled, blurred, and noisy images by accurate total variation minimization with local constraints. *Multiscale Modeling & Simulation*, 5(1):235–272, 2006.
- [5] M. Arigovindan, M. Sühling, P. Hunziker, and M. Unser. Variational Image Reconstruction from Arbitrarily Spaced Samples: A Fast Multiresolution Spline Solution. *IEEE Transactions on Image Processing*, 14(4):450–460, 2005.
- [6] R. Bass and K. Gröchenig. Random sampling of multivariate trigonometric polynomials. *SIAM Journal of Mathematical Analysis*, 36:795, 2004.
- [7] A. Beck and M. Teboulle. Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems. *IEEE Transactions on Image Processing*, 18:2419–2434, November 2009.
- [8] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Img. Sci.*, 2:183–202, March 2009.
- [9] Maïtine Bergounioux and Loic Piffet. A second-order model for image denoising. *Set-Valued and Variational Analysis*, 18:277–306.
- [10] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, 2003.
- [11] S. Bigot-Marchand. *Outils de traitement d’images adaptés au traitement d’images omnidirectionnelles*. PhD thesis, Oct 2008.
- [12] J. M. Bioucas-Dias and M. A. T. Figueiredo. A New TwIST: Two-Step Iterative Shrinkage/Thresholding Algorithms for Image Restoration. *IEEE Transactions on Image Processing*, 16:2992–3004, December 2007.
- [13] T. Blumensath and M. Davies. Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, 14:629–654, 2008. 10.1007/s00041-008-9035-z.

- [14] T. Blumensath and M. E. Davies. On the difference between orthogonal matching pursuit and orthogonal least squares. 2007.
- [15] D. Blumstein et al. . IASI instrument: technical overview and measured performances. In M. Strojnik, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5543 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 196–207, November 2004.
- [16] F. Brachet et al. . Static Fourier transform spectroscopy breadboards for atmospheric chemistry and climate. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7100 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, September 2008.
- [17] E. Breton, A. Brouillon, R. Gachet, and F. De Lussy. Preflight and in-flight geometric calibration of SPOT5 HRG and HRS images. *ISPRS Comm. I, Denver, CO, 10-15 Nov 2002*.
- [18] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. 2004.
- [19] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.
- [20] E. Cansot et al. . Static Infrared Fourier Transform Interferometer (SIFTI): Benefits of Phase Modulation Processing. In *Fourier Transform Spectroscopy/ Hyperspectral Imaging and Sounding of the Environment*, page FMC3. Optical Society of America, 2007.
- [21] V. Caselles, A. Chambolle, and M. Novaga. The discontinuity set of solutions of the tv denoising problem and some extensions. *Multiscale Modeling & Simulation*, 6(3):879–894, 2007.
- [22] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20:89–97, 2004.
- [23] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997. 10.1007/s002110050258.
- [24] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40:120–145, 2011.
- [25] S. Chen, B. A. Billings, and W. Luo. Orthogonal Least Squares Methods and Their Application to Non-linear System Identification. *International Journal of Control*, 50(5):1873–1896, 1989.
- [26] S. Chen, D. L. Donoho, and M. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM J. Sci. Comput.*, 20:33–61, December 1998.
- [27] C. K. Chui and X. Shi. Bessel Sequences and Affine Frames. *Applied and Computational Harmonic Analysis*, 1:29–49, Dec 1993.
- [28] P.-L. Combettes. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, 53 n° 5-6:475–504, 2004.
- [29] P.-L. Combettes and Bauschke H. H. *Convex analysis and Monotone Operator Theory in Hilbert Spaces*. Springer Science, 2011.

- [30] P.-L. Combettes and V.R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Journal on Multiscale Modeling and Simulation*, 4 n° 4:1164–1200, 2005.
- [31] B. Coulange. *Detection et correction de l'aliasing par extension du signal analytique aux images numériques*. PhD thesis, Nov. 2011. Thèse de doctorat.
- [32] A. J. Coulson. A generalization of nonuniform bandpass sampling. *IEEE Transactions on Signal Processing*, 43:694–704, March 1995.
- [33] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, 2004.
- [34] G. M. Davis, S. Mallat, and Z. Zhang. Adaptive time-frequency decompositions. *Optical Engineering*, 33:2183–2191, July 1994.
- [35] Julie Delon and Bernard Rougé. Small baseline stereovision. *Journal of Mathematical Imaging and Vision*, 28:209–223, 2007.
- [36] A. Desolneux, L. Moisan, and J.-M. Morel. Meaningful alignments. *International Journal of Computer Vision*, 40:7–23, 2000.
- [37] A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis: A Probabilistic Approach*. Springer-Verlag, collection "Interdisciplinary Applied Mathematics", 2008.
- [38] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52:1289–1306, 2006.
- [39] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck. Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit. *Discovery*, (March):1–39, 2006.
- [40] Jean Duchon. Splines minimizing rotation-invariant semi-norms in sobolev spaces. In Walter Schempp and Karl Zeller, editors, *Constructive Theory of Functions of Several Variables*, volume 571 of *Lecture Notes in Mathematics*, pages 85–100. Springer Berlin / Heidelberg, 1977.
- [41] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Transactions of the American Mathematical Society*, 72(2):341–366, Mar 1952.
- [42] S. Durand, F. Malgouyres, and B. Rougé. Image deblurring, spectrum interpolation and application to satellite imaging. *ESAIM: Control, Optimisation and Calculus of Variations*, (5):445–475, 2000.
- [43] M. Elad. Why simple shrinkage is still relevant for redundant representations. *IEEE Transactions on Information Theory*, 52:5559–5569, 2006.
- [44] M. Elad, B. Matalon, J. Shtok, and M. Zibulevsky. A wide-angle view at iterated shrinkage algorithms. In *in SPIE Wavelet XII*, pages 26–29, 2007.
- [45] L. Gross-Colzy F. de Lussy, D. Greslou. Process line for geometrical image correction of disruptive microvibrations.
- [46] G. Facciolo, A. Almansa, J.-F. Aujol, and V. Caselles. Irregular to regular sampling, denoising, and deconvolution. *Multiscale Modeling & Simulation*, 7(4):1574–1608, 2009.

- [47] S. J. Favier and R. A. Zalik. On the stability of frames and Riesz bases. *Applied and Computational Harmonic Analysis*, 2:160–173, 1995.
- [48] H. G. Feichtinger, K. Gröchenig, and T. Strohmer. Efficient numerical methods in non-uniform sampling theory. *Numerische Mathematik*, 69:423–440, 1995. 10.1007/s002110050101.
- [49] Hans G. Feichtinger and Karlheinz Gröchenig. Iterative reconstruction of multivariate band-limited functions from irregular sampling values. *SIAM J. Math. Anal.*, 23:244–261, January 1992.
- [50] K. Grochenig. Acceleration of the frame algorithm. *IEEE Transactions on Signal Processing*, 41(12):3331–3340, December 1993.
- [51] P. Hébert et al. . Instrumental Concept and Preliminary Performances of SIFTI: Static Infrared Fourier Transform Interferometer. International Conference on Space Optics, Toulouse, 2008.
- [52] C. Herley and P. W. Wong. Minimum rate sampling and reconstruction of signals with arbitrary frequency support. *IEEE Transactions on Information Theory*, 45:1555–1564, July 1999.
- [53] M. I. Kadec. The exact value of the Paley-Wiener constant. *Sov. Math. Dokl.*, 5:559–561, 1964.
- [54] J. Kalifa, S. Mallat, and B. Rougé. Deconvolution by thresholding in mirror wavelet bases. *IEEE Transactions on Image Processing*, 12:446–457, April 2003.
- [55] Y. Katznelson. *An introduction to Harmonic Analysis*. Cambridge Mathematical Library, Third corrected edition, 2004.
- [56] J. Keiner, S. Kunis, and D. Potts. Using nfft 3—a software library for various nonequispaced fast fourier transforms. *ACM Trans. Math. Softw.*, 36:19:1–19:30, August 2009.
- [57] A. Kohlenberg. Exact interpolation of band-limited functions. *Journal of Applied Physics*, 24:1432–1436, December 1953.
- [58] H. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Mathematica*, 117:37–52, 1967.
- [59] L. Lebègue, Ph. Kubik, D. Greslou, F. Delussy, and N. Theret. Using exotic guidance for PLEIADES-HR image quality calibration. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (XXXVII), 2008.
- [60] N. Levinson. *Gap and Density Theorems*, volume 26. American Mathematical Society Colloquium Publications, 1940.
- [61] P. Lier, C. Valorge, and X. Briottet. *Imagerie Spatiale*. 2008.
- [62] Yuan-Pei Lin and P. P. Vaidyanathan. Periodically Nonuniform Sampling of bandpass Signals. *IEEE Transactions on Circuits and Systems*, 45(3):340–351, March 1998.
- [63] J. G. Liu and G. L. K. Morgan. FFT Selective and Adaptive Filtering for Removal of Systematic Noise in ETM+ Imageodesy Images. *IEEE Transactions on Geoscience and Remote Sensing*, 44:3716–3724, December 2006.

- [64] F. Malgouyres. Combining total variation and wavelet packet approaches for image deblurring. In *Proceedings of the IEEE Workshop on Variational and Level Set Methods (VLSM'01)*, VLSM '01, pages 57–, Washington, DC, USA, 2001. IEEE Computer Society.
- [65] S. Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd edition, 2008.
- [66] Y. Meyer. Du nouveau sur l'échantillonnage ou comment tricher avec Shannon.
- [67] Lionel Moisan. How to discretize the total variation of an image? *PAMM*, 7(1):1041907–1041908, 2007.
- [68] Tikhonov A. N. and Arsenin V. Y. *Solutions of ill posed problems*. Wiley, New York, 1977.
- [69] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.*, 103(1):127–152, 2005.
- [70] Y. Nesterov. Gradient methods for minimizing composite objective function. CORE Discussion Papers 2007076, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), September 2007.
- [71] N. Nguyen, P. Milanfar, and G. Golub. A computationally efficient superresolution image reconstruction algorithm. *IEEE Transactions on Image Processing*, 10:573–583, April 2001.
- [72] Weiss P. Algorithmes rapides d'optimisation convexe, applications à la restauration d'images et à la détection de changements, 2008.
- [73] R. E. A. C. Paley and N. Wiener. Fourier transforms in the complex domain. *American Mathematical Society Colloquium Publications*, 19, 1934.
- [74] C. Pierangelo et al. . SIFTI, a Static Infrared Fourier Transform Interferometer dedicated to ozone and CO pollution monitoring. Presented at the 16th International TOVS Study Conference, May 2008.
- [75] G. Rochefort, F. Champagnat, G. Le Besnerais, and J.-F. Giovannelli. An Improved Observation Model for Super-Resolution Under Affine Motion. *IEEE Transactions on Image Processing*, 15:3325–3337, November 2006.
- [76] R. T. Rockafellar. *Convex Analysis*. Princeton Univ. Press, 1970.
- [77] C. D. Rodgers. *Inverse Methods for Atmospheric Sounding, Theory and Practice*, volume 2 of *Series on Atmospheric Oceanic and Planetary Physics*. World Scientific.
- [78] S. Roques, F. Brachere, B. Rougé, and M. Pausader. Séparation des décalages induits par l'attitude et le relief entre images d'un couple stéréoscopique. 18<sup>e</sup> *Colloque sur le traitement du signal et des images*, pages 765–768, 2001.
- [79] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259 – 268, 1992.
- [80] L. Moisan S. Durand. Correction des perturbations par microvibrations des nappes de disparités sans modèle de microvibration a priori. *Rapport de R&T CNES*, December 2010.
- [81] N. Sabater. Fiabilité et précision en stéréoscopie application à l'imagerie aérienne, 2009.

- [82] Neus Sabater, Andrés Almansa, and Jean-Michel Morel. Rejecting wrong matches in stereovision. *CMLA preprint*, July 2008.
- [83] Neus Sabater, Andres Almansa, and Jean-Michel Morel. Meaningful matches in stereovision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99(PrePrints), 2011.
- [84] D. Simeoni et al. . Design and development of IASI instrument . In M. Strojnik, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5543 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 208,219, November 2004.
- [85] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V - The discrete case. *AT T Technical Journal*, 57:1371–1430, June 1978.
- [86] M. V. Solodov. A class of decomposition methods for convex optimization and monotone variational inclusions via the hybrid inexact proximal point framework. *Optimization Methods and Software*, 19(5):557–575, 2004.
- [87] T. Strohmer. Computationally attractive reconstruction of bandlimited images from irregular samples. *IEEE Transactions on Image Processing*, 6:540–548, April 1997.
- [88] Paul Tseng and Dimitri P. Bertsekas. On the convergence of the exponential multiplier method for convex programming. *Mathematical Programming*, 60:1–19, 1993.
- [89] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing: Part I-theory. *IEEE Trans. Signal Processing*, 41:821–833, 1993.
- [90] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing II efficiency design and applications. *Signal Processing, IEEE Transactions on*, 41(2):834–848, August 2002.
- [91] P. Vandewalle, S. E. Süsstrunk, and M. Vetterli. Superresolution images reconstructed from aliased images. volume 5150 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1398–1405, June 2003.
- [92] R. G. Vaughan, N. L. Scott, and D. R. White. The theory of bandpass sampling. *IEEE Transactions on Signal Processing*, 39:1973–1984, September 1991.
- [93] C. Vazquez, E. Dubois, and J. Konrad. Reconstruction of nonuniformly sampled images in spline spaces. *IEEE Transactions on Image Processing*, 14:713–725, June 2005.
- [94] J. J. Wojtiuk and R. J. Martin. Random sampling enables flexible design for multiband carrier signals. *IEEE Transactions on Signal Processing*, 49:2438–2440, October 2001.
- [95] R.M. Young. An introduction to nonharmonic Fourier series. 1980.



## Résumé

Les performances des instruments d'acquisition satellitaire progressent rapidement grâce au développement des technologies mais aussi grâce à la compréhension et l'intégration des phénomènes physiques complexes intervenant lors de l'acquisition. Cette thèse traite de plusieurs problèmes d'échantillonnage irrégulier dont les micro-vibrations des satellites dits push-broom tels que SPOT5 et les récents satellites Pléiades dont les capacités en imagerie permettent la détermination de modèles d'élévation très précis. Nous traitons aussi de l'inversion d'interférogrammes en spectrogrammétrie où l'irrégularité de l'échantillonnage est liée à la précision d'usinage des composants réfléchissants. Les microvibrations dans le cas du tangage sont estimées à partir d'une nappe de disparité altérée et non-dense par contraintes de parcimonie. Nous montrons expérimentalement que ce modèle et les algorithmes utilisés permettent de résoudre en partie ce problème mal posé. L'ajout d'un a priori sur la régularité de l'élévation permet d'améliorer encore cette estimation dans les cas plus difficiles. Les images acquises en présence de microvibrations nécessitent de plus un rééchantillonnage auquel s'ajoute la déconvolution avec une problématique de coût numérique. L'algorithme que nous présentons ici répond à ces besoins grâce au cadre fonctionnel des splines que nous adaptons au problème de la déconvolution, avec des performances équivalentes à l'état de l'art et un coût numérique maîtrisé. Enfin nous abordons un problème inverse en interférométrie statique où la nature des signaux et de l'échantillonnage soulève de nombreuses questions, ce travail réalisé lors d'une R&T sur l'instrument SIFTI développé au CNES y apporte des réponses claires sous forme de résultats théoriques et numériques dans le cadre unifié des séries de Fourier non-harmoniques.

## Abstract

Performances of remote-sensing satellite instruments have been increasing fast thanks to developing technologies but also thanks to the better understanding and integration of complicated physical phenomena occurring during the acquisition. This thesis report addresses several irregular sampling problems including the microvibrations of push-broom satellites such as SPOT5 and recently launched Pléiades satellites whose imaging capabilities allow for the computation of very accurate numerical elevation models. We also address the inversion of interferograms in spectrogrammetry where the sampling irregularity comes from the imperfectly machined reflecting components. Microvibrations in the single pitch case are estimated from a perturbed, non-dense disparity map with sparsity constraints. Experiments show that this modelling and the developed algorithms enable partial resolution of this ill-posed problem. Furthermore, an additional regularity hypothesis on the elevation improves this estimation for more difficult cases. Images suffering from micro-vibrations during acquisition necessitate a sampling correction together with deblurring and fast restoration. The algorithm we present here fulfils these requirements thanks an adaptation of the splines setting to the deblurring case, it is faster than state-of-the-art algorithms with equivalent performances. Finally we address the interferogram inversion problem in which the signals and sampling sets raise many questions, this work was achieved during a R&T study of the SIFTI instrument examined at CNES, it clarifies these questions in the form of theoretical and numerical results using the unifying framework of nonharmonic Fourier series.